FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Real-time detection of FOG episodes in patients with Parkinson's Disease

Susana Perdigão de Sousa

DISSERTATION FOR THE DEGREE OF MASTER IN BIOENGINEERING

MESTRADO INTEGRADO EM BIOENGENHARIA - ENGENHARIA BIOMÉDICA

Supervisor at FEUP:: Aníbal Ferreira, PHD

Supervisor at FhP-AICOS:: Vânia Guimarães, MSC

July 8, 2018

Real-time detection of FOG episodes in patients with Parkinson's Disease

Susana Perdigão de Sousa

MESTRADO INTEGRADO EM BIOENGENHARIA - ENGENHARIA BIOMÉDICA

July 8, 2018

Resumo

A doença de Parkinson é uma doença neurodegenerativa progressiva que afecta 1% da população mundial com mais de 65 anos. Dado que se trata de uma doença sem cura conhecida, o tratamento farmacológico tenta apenas atenuar os efeitos dos sintomas associados. Os mais comuns e perturbadores pertencem ao domínio motor. Um sintoma típico nos estágios mais avançados é o Congelamento da Marcha - *Freezing of Gait* - que consiste num súbito bloqueio do movimento, descrito pelos pacientes como que os seus pés ficassem presos ao chão. Desta forma, é uma das principais causas de quedas nestes pacientes afectando significativamente a sua qualidade de vida.

Actualmente, a avaliação dos sintomas é feita a partir da aplicação de questionários e escalas de avaliação, que levam geralmente a uma avaliação errada do estado do paciente e da evolução da doença. A quantificação e monitorização objectiva dos episódios de *freezing*, através da utilização de tecnologias *wearable*, pode ter um papel importante com vantagens tanto para os pacientes como para os especialistas. A combinação da monitorização destes eventos com estratégias de estimulação auditiva, possibilita a criação de um sistema capaz de ajudar os médicos a gerir os tratamentos e, por outro lado, ajudar os pacientes a lidar e a superar os eventos de *freezing*.

Tendo isto em mente, no âmbito desta dissertação desenvolveu-se uma metodologia para monitorização e detecção dos episódios de *Freezing of Gait*. Trata-se de um algoritmo de aprendizagem supervisionada, no qual se usam sinais de aceleração, adquiridos por sensores inerciais colocados no tornozelo do paciente, como entrada de um classificador - *Support Vector Machines*. Como os episódios de *freezing* têm uma natureza muito dependente do paciente a metodologia proposta é uma metodologia personalizada. Além de utilizar apenas dados do paciente em questão para a construção do modelo de classificação, esta abordagem inclui a segmentação do sinal com base no ciclo de marcha de cada indivíduo e a extracção de medidas de similaridade entre um modelo da marcha e as porções do sinal de aceleração segmentado, através do cálculo de *Dynamic Time Warping*. Características de frequência são também propostas. A metodologia proposta foi avaliada com recurso a dados de aceleração de 8 pacientes presentes no dataset DAPHNet, obtendo-se 82,0% de especificidade, 85,1% de sensibilidade e 58,9% de *F1-score*. Comparando as características propostas com um conjunto de características da literatura, os desempenhos foram muito similares, com uma melhoria nos resultados quando os conjuntos são usados simultaneamente.

Adicionalmente, realizou-se uma análise preliminar com vista à previsão de eventos de *Freezing of Gait* antes do seu início. Para tal, comparou-se o comportamento das características propostas durante períodos de marcha normal e imediatamente antes dos eventos de bloqueio da marcha. Os resultados sugerem que a fase anterior a estes eventos é também muito dependente do paciente em questão e que existem características capazes de capturar a degradação da marcha antes do início de um *Freezing of Gait*.

A principal limitação dos resultados é o elevado número de falsos positivos detectados. No entanto, os resultados foram promissores e sugerem que a metodologia proposta pode ser usada para detectar eventos de *freezing* em pacientes com doença de Parkinson. No futuro, novos desenvolvimentos podem ser necessários para validar e melhorar os resultados.

Abstract

Parkinson's Disease is a progressive neurodegenerative disorder that affects 1% of the worldwide population over the age of 65. Being a disease without a known cure, the pharmacological treatment tries to mitigate the effects of associated symptoms. The most common and disturbing ones belong to the motor domain. A typical impairment, in the advanced stages, is the Freezing of Gait. It represents a sudden blocking of movement described by the patients as their feet are stuck to the ground. Hence, this is the main cause of falls and significantly affects patients quality of life.

Currently symptoms evaluation is performed through the application of questionnaires and rating scales, which usually lead to erroneous assessment of the patient state and disease evolution. Objective quantification and monitoring of freezing episodes, through the use of wearable technologies, would have a key role, with advantages for both patients and clinicians. Combining freezing events monitoring to auditory stimulation strategies, one is able to build a system capable of helping clinicians to manage disease treatments and, on the other side, of helping patients to deal and overcome these events.

Having this is mind, a methodology for monitoring and detecting Freezing of Gait episodes was developed and is presented in this dissertation. It is a supervised machine learning algorithm, where acceleration signals acquired from inertial sensors placed in the patient's ankle are used as the input of a supervised machine learning classifier - Support Vector Machines. Due to the patient dependent nature of the freezing events, a personalized methodology is proposed. In addition to only using data from the concerned patient, the patient-dependent approach includes the segmentation of the signal based on the gait cycle of each individual and also the extraction of similarity measures between a gait template and portions of the recorded acceleration signal, through the calculation of the Dynamic Time Warping. Frequency-based features are also proposed. The methodology was evaluated with acceleration data of 8 patients from the DAPHNet dataset, achieving a specificity of 82,0%, a sensitivity of 85,1% and a F1-score of 58,9%. The proposed features were compared with a set of features proposed in the literature and the performances were very similar, with an improvement in the results when the two sets are used together.

Additionally, a preliminary analysis towards prediction of freezing events before their beginning was also performed by comparing the behavior of the proposed features during normal gait and prior to the freezing episodes. The results suggest that the pre freezing phase is also highly patient dependent and that there are features able to capture gait degradation prior to Freezing of Gait.

The main limitation of the results is the high number of false positives detected. However, the results were promising and suggest that the proposed methodology can be used to detect freezing events in patients with Parkinson's Disease. Further developments to this project are required to better validate and also improve the results.

Acknowledgments

Em primeiro lugar, gostaria de agradecer ao professor Aníbal Ferreira pela disponibilidade, pela experiência e apoio prestado ao longo deste projecto. À minha orientadora Vânia Guimarães, por todo o acompanhamento, motivação e apoio, pelo conhecimento e por todas as sugestões que permitiram ultrapassar os pequenos obstáculos e progredir no projecto. A toda a Fraunhofer AICOS Portugal pelas excelentes condições de trabalho fornecidas.

Quero também agradecer aos meus amigos. Aos que estiveram sempre lá durante estes 5 anos, tornando este percurso memóravel. Aos que me ajudaram sempre que precisei, tanto em questões mais complexas como apenas a descomplicar o que era simples. Àqueles com quem partilhei frustrações, alegrias e sucessos. Àqueles que se juntavam ao fim de um dia de trabalho só para descontrair e perceber que estavamos todos no mesmo barco. Ao Pedro. Pelas palavras sábias nos momentos certos. Pela força e por acreditar mais em mim do que eu própria, muitas vezes. Pelo abraço reconfortante no fim de cada dia.

E o mais especial, o obrigada à minha família. Ao meu irmão por me aturar todos os dias, pela paciência e pelos sábios conselhos. Aos meus pais, por estarem sempre presentes, pelo apoio incondicional, por permitirem que tudo isto fosse possível. E a toda a restante família, mesmo aqueles que sei que estarão algures muito orgulhosos por ter concluído mais esta etapa.

Susana Sousa

"We'll try again and we'll fail again because that's what progress looks like.
Progress looks like a bunch of failures.
And you're going to have feelings about that because it's sad, but you cannot fall apart.
And then one day, we will succeed..."

Meredith Grey, *Grey's Anatomy*

Contents

List of Figures

List of Tables

List of Abbreviations

ADL	Activities of Daily Living
ANN	Artificial Neural Networks
DBA	DTW Barycenter Averaging
DFT	Discrete Fourier Transform
DTW	Dynamic Time Warping
FFT	Fast Fourier Transform
FI	Freezing Index
FOG	Freezing of Gait
GM	Geometric Mean
H&Y	Hoehn & Yahr
IMU	Inertial Measurement Units
MB	Moore and Bachlin method
MDS-UPDRS	Movement Disorder Society - Unified Parkinson's Disease Rating Scale
MEMS	Micro-Eletro-Mechanized-Systems
MI	Mutual Information
PD	Parkinson's Disease
PI	Power Index
QOL	Quality of Life
REM	Rapid Eye Movement
RF	Random Forests
SFS	Sequential Forward Selection
SVM	Support Vector Machines

Chapter 1

Introduction

The technological developments in wearable sensors systems is giving rise to a new concept of healthcare. These systems are being integrated in a huge number of clinical applications, from diagnosis to quantitative and objective monitoring of human activities [1, 2].

One of the main applications of wearable sensors is in the monitoring of movement, relying on inertial systems, for rehabilitation and assistance of unhealthy and elderly people [3], on both clinical and home environments [4, 5]. It has a major importance in the study and detection of falls [6], in rehabilitation interventions in stroke survivors [2], to evaluate movement in patients with Alzheimer's Disease [7] and in the assessment of movement disorders associated with Parkinson's Disease (PD) [6].

The continuous monitoring of these patients is fundamental in helping clinicians to objectively quantify the symptoms over time and evaluate the motor status of an individual. The follow-up process is facilitated, as doctors can access more data than that collected during the typical short lasting appointments [2]. Thus, spatial and time barriers are eliminated. Moreover, these systems can be developed with feedback features [8] that not only inform patients about their status, but also intend to reduce the burden of the caregiver by alerting when there is a change in patient condition or in case a medical decision has to be made [2] - remote assistance.

Since PD is a complex neurodegenerative disorder, with the most common symptoms involving the motor domain, these technologies have here a high relevance. Among others, PD symptoms include muscular rigidity, tremors, postural instability, bradykinesia and episodes of Freezing of Gait (FOG) [4, 9]. The latter represents one of the most disabling conditions, especially in elderly long-term, advanced patients, who describe the events as their "feet are glued to the ground". Thus, this is the main cause of falls in patients suffering from PD [10].

The clinical diagnosis and evaluation of PD is mainly performed through visual examination of patient motor performance. Therefore, an objective quantification of the patient motion would lead to an accurate diagnosis, therapy efficacy assessment and a better follow-up of the disease's evolution, overtaking issues associated to subjectivity and inter-professional variability [4].

In addition, motor disabilities directly affect patient's quality of life (QOL). So, it is important not only to monitor the disease, but also to reduce the burden of its symptoms, as is the case of

FOG events. Studies demonstrated that visual and auditory stimulus are able to improve these motor symptoms [11, 12], improving the gait stability of the patients. Wearable technologies can integrate such features in their systems and have an active role when the patient undergoes a FOG event. Thereby, it is possible to prevent or reduce the effect of such symptoms, improving the gait quality of patients, and consequently their independence.

1.1 Problem identification

Nowadays, the gold standard for clinical assessment of PD symptoms is the application of clinical rating scales [4, 13], some of them specifically designed for FOG evaluation [14]. Although validated, they can be unreliable due to the subjective perspective of patients and caregivers and also to the subjective analysis by the clinician. Most of the time, self-assessment is also inappropriate due to cognitive impairments or memory loss of the patients [15]. Hence, symptoms quantification for both diagnosis and disease evaluation is compromised.

Video recorded gait is also often used for FOG evaluation, since it allows the identification of FOG episodes by different professionals in a *post-hoc* analysis [16] and is the gold standard for FOG assessment [4]. Other technologies based on movement measurement systems (ultrasound and optical systems and pressure platforms) have been used, although they are not well accepted because they are expensive, unportable and only adequate for laboratory environments [4, 5, 17].

Wearable technologies based on inertial systems have been investigated. Several studies have used inertial sensors to improve the analysis and detection of FOG events, through its placement on different segments of the body [8, 17, 18, 19, 20]. They tried to quantify the characteristics of FOG events and to implement systems for real time FOG detection with rhythmic auditory cueing to prevent or facilitate the overcoming these episodes.

However, there is still a lack of consensus regarding the optimal number of sensors and the optimal site for their placement as well as the most effective signal processing algorithms [4, 6]. Furthermore, most of the studies only focus on the detection of FOG, requiring a minimum latency from their start, and thus the auditory stimulation is non effective in the prevention.

Therefore, the problem relies on the need of an effective way of detecting FOG episodes with low latency, or even predict their occurrence, through the continuous monitoring of the patient movement using low cost wearable systems, in a comfortable and easy to use configuration, allowing auditory stimulation and increasing the patient QOL.

1.2 Motivation and objectives

There is a wide variety of techniques and sensors that have already been explored to assess motor symptoms of PD. Some of them are only used in the clinical context and others are also adequate to home environments.

Taking into account the problem identified above and drawbacks of current solutions, inertial sensors systems seem to be the most appropriate way to measure motor performance of the patients. This kind of systems is usually well accepted by the patients since they are easy to use and comfortable to be worn, due to their small size and weight. They are portable, low-cost and sufficiently accurate in the measurements. They are not restricted to a clinical environment, being also used in unsupervised environments, allowing the monitoring and assistance of PD patients during common activities of daily living (ADL) [4, 5, 8, 17].

Thus, these systems can be used for an automatic detection and objective quantification of Parkinson's disease motor symptoms, particularly FOG episodes. They allow the assessment of patient motor status and enable the evaluation of treatment's efficacy, which has the potential to improve medication management. In addition, such a system can reduce the number of medical interventions, reducing associated costs and empowering the patient to play an active role in the management of the disease. Since these episodes have a daily, unpredictable and frequent occurrence, they have a significant impact on fall risk and QOL [21], so patients would benefit a lot with this kind of technology.

In this context, the main objective of this dissertation is the development of a solution for motor symptoms detection in PD, namely, the detection of FOG episodes and pre-FOG patterns. Moreover, data acquisition with PD patients, using inertial sensing devices placed on distinct body positions, and the development of a proof of concept are additional goals of this work.

1.3 Document structure

This document reports all the work performed in this dissertation and it is organized in chapters and sub-chapters:

Chapter 1 presents the Introduction, where the first approach to the theme of the dissertation is made, identifying the gap that this work tries to fill and then the main goals.

Chapter 2, named as Background and Literature Review, provides a detailed overview of the main aspects related to this work. It starts with a description of the fundamentals behind Parkinson's Disease and its symptoms, followed by a brief overview of the type of sensors that will be used and then a literature review of the main studies developed in this field.

In Chapter 3, Methodology for Detecting Freezing of Gait episodes, the development steps of the proposed methodology are presented along with the theoretical fundamentals needed to understand each step. A detailed description and understanding of the public dataset used during this entire work is also performed.

Chapter 4 presents the methodology for a preliminary analysis of the pre-FOG patterns, adopted in this work.

In Chapter 5, named as Results and Discussion, a detailed analysis of the results obtained in the main steps of the proposed methodology is performed along with their discussion.

Chapter 6, presents the Conclusions and the Future Work that could be done to improve the results of the proposed methodology.

Chapter 2

Background and literature review

This chapter presents an overview of the existing literature on the fundamental topics regarding monitoring of freezing episodes in patients with Parkinson's Disease.

First of all, to better understand the impact of the disease and its symptoms and the existing medical strategies, Section 2.1 provides an overview of the main clinical aspects related to PD, with special attention to FOG, since it will be the main focus of this work.

Then, in Section 2.2, a detailed explanation is presented about Inertial Sensor Systems that will have a critical role on this work. Here, their advantages and the main associated technological aspects are described. These systems have gained relevance in the activity monitoring field and thus, over the years, several studies have been performed aiming to monitor FOG symptoms through the use of wearable inertial systems. Section 2.3 provides a summary of the main works developed in the field, briefly describing the proposed methodology and the main innovative aspects of each one. To better understand the characteristics of normal gait and the correspondent inertial signals, Section 2.4 performs a briefly gait analysis.

2.1 An Overview of Parkinson's Disease

Parkinson's Disease is a complex neurodegenerative disorder which affects approximately 1% of the worldwide population over the age of 65 [22].

It is a chronic and progressive disease mainly characterized by symptoms involving the motor domain, such as bradykinesia, resting tremor, postural instability and rigidity, which are usually considered as the four major motor signs of PD [9]. Over time, a large variety of late onset motor symptoms appear, typically freezing of gait, flexed posture, speech difficulties and others. Non-motor features are also associated to PD, including olfactory disturbances, autonomic dysfunction, sleep fragmentation and depression [9, 23].

In the following sections an overview of the main aspects related to PD is presented. Section 2.1.1 presents a brief explanation of the physiological mechanisms behind PD. Then, in Section 2.1.2, the main motor and non-motor associated symptoms are described, with special attention to FOG, since it is the focus of this dissertation. The main strategies for clinical assessment of PD

and particularly of FOG are analyzed in Section 2.1.3 and, finally, Section 2.1.4 presents a simple overview on the typical treatment.

2.1.1 Pathophysiology of Parkinson's Disease

It is known that the pathophysiology of PD is related to the concentration of dopamine in the brain. Dopamine is a neurotransmitter produced by neurons of the substantia nigra - a structure located in the mid brain. It has an important role in movement control. When a person suffers from PD, the dopaminergic neurons progressively degenerate leading to a significant decrease of dopamine available in the brain and consequently to the appearance of motor deficits [24].

Although the physiologic mechanisms that lead to motor impairments are well understood, the knowledge and information about PD is still limited and the causes behind progressive degeneration of dopaminergic neurons remain poorly understood [25]. Therefore, there is a lack of understanding of the associated risk factors and thus, a lack of effective therapies to prevent or improve both motor and non-motor symptoms [22].

2.1.2 Parkinson's Disease Symptoms

The manifestation of the symptoms varies with the disease evolution, drug therapy, environmental and emotional factors and is also patient dependent [20].

Although some research indicates the existence of a pre-motor phase, where the individuals present early non-motor signs of PD, the current clinical diagnosis is performed through the assessment of motor symptoms [23]. During the first stages of the disease, motor abnormalities may be minimal and thus, most of the patients remain untreated until motor features become apparent and disabling, affecting patient's daily life activities. As the disease evolves, the severity of both motor and non-motor symptoms increases, compromising patient activities and even leading to loss of independence [22]. Table 2.2 summarizes the symptoms of PD and the more common ones are described in the next sections.

2.1.2.1 Motor Symptoms

As mentioned before, the main motor symptoms of PD are tremor, postural instability, rigidity and bradykinesia. In advanced stages, gait impairments, such as FOG, become relevant. The worsening of these symptoms can be very disabling for PD patients, as one can see below.

Tremor namely resting tremor, is the most common symptom of PD, appearing in almost 70% of patients [4, 9]. It is an easily recognized symptom, typically affecting one side of the body, in the distal parts of the limbs, such as fingers and hands, evolving to the jaw, feet and legs, in later stages. It is present when the patient is at rest or when the limb is held in one position, disappearing with motion [9, 26]. There are other subtypes of tremor, though they are associated with different movement disorders, they interfere with the diagnostic process of PD.

Rigidity is characterized by the increase in tightness or stiffness of the body muscles. It is present during the passive movement of a limb and can be manifested along with pain [9]. For this reason, sometimes it is wrongly associated with arthritis or orthopedic problems. In the later stages of the disease, postural deformities may also occur, that are associated with the increase of rigidity of the neck, trunk, elbows and knees. These deformities can include neck flexion, truncal flexion and scoliosis [27].

Bradykinesia is an easily recognized symptom of PD, referring to slowness of movement [9]. The patient with bradykinesia has difficulties in preparing the instructions to move, in the onset and during the execution of movements, since the muscles have not enough energy to initiate and perform movement. Reaction times are also compromised, increasing with the advance of the disease [28]. Bradykinesia is manifested during the performance of ADL, from the ones that require simple movements, to the more complex ones, that need simultaneous, sequential or repetitive movements [9, 28], such as writing, brushing, buttoning, clapping and using utensils. It is important to notice that the term bradykinesia is often used along with the terms akinesia and hypokinesia. Although very similar, these concepts present slight differences: akinesia refers to the absence of spontaneous movement while hypokinesia refers to the decreased amplitude of movement. As stated, bradykinesia is mostly related with the slowness of movement, but the term is frequently used to express the three impairments [29].

Postural instability is one of the later manifestations of PD and it appears due to the loss of postural reflexes [9]. The patients are not capable of maintaining a steady and a upright posture and thus have a high risk of falling [30].

Freezing of gait As mentioned before, FOG is the main focus of this dissertation. Usually present at the advanced stages of the disease, FOG is a common and one of the most disabling symptoms of PD [31]. Based on a survey performed with 6620 patients of the German Parkinson Association [32], 47% of the patients report experiencing FOG regularly, 28% of them with a daily frequency. Among severely affected patients, it affects about 80% of them. It is also more likely in men than in women. FOG is a complex and heterogeneous phenomenon defined as an episodic inability to produce effective stepping, described by the patients as if their feet were "glued to the ground" [33].

These episodes are context - dependent and, according to the subgroup classification of Fahn [34], there are five types of FOG, depending on the situations where it mainly appears:

- start hesitation - when the patient initiates gait process;
- turn hesitation - while the patient tries to make a turn, specially in narrow turns;
- hesitation in tight quarters - when the patient approaches narrow spaces, as doors or corridors;

- destination hesitation - when the feet freeze as the patient approaches a target;
- open space hesitation - a spontaneous freezing without an apparent provoking factor.

In addition, FOG is more common in patients home, during unobserved behavior, than in clinical environment [35]. These episodes last from seconds to minutes, which is highly associated with the disease stage [31]. First, patients steps start shortening (pre-freezing state) [36] and then he/she stops (freezing state). During this period, the patient attempts to make steps that may lead to a loss of postural balance and provoke falls and injuries. Moreover, FOG compromises ADL and social contacts, contributing to the social burden of the disease and a decrease in QOL [32].

The mechanisms responsible for this phenomenon are not yet clarified, but the literature suggests 4 models that try to explain the potential mechanisms behind the episodic nature of FOG. These models are explained below and are summarized in Table 2.1.

- **Threshold model** - proposed by Plotnik M. *et al.* [37], this model states that the various gait impairments associated with FOG can interact to a point of motor breakdown, leading the locomotion system to a freezing event.
- **Interference model** - Lewis and Barker [38] suggested that FOG occurs when the patient faces an increasing number of concurrent and challenging tasks, activating both the cognitive and the motor systems. The occurrence of a momentary breakdown of information processing leads to an interruption of gait.
- **Cognitive model** - this model, proposed by Vandenbossche *et al.* [39], associates FOG to situations where a decision has to be made. Situations that require a response decision involve both automatic and consciously controlled mechanisms. These are more severely impaired in patients that manifest freezing events, and thus these systems are more subjected to breakdown, leading to behavioral indecision;
- **Decoupling model** - proposed by Jacobs *et al.* [40], suggests that FOG is an inability to combine the various phases of step initiation - anticipatory postural adjustments. It is a disconnection between preparatory programming and the intended motor response.

Some studies associate the appearance and severity of FOG with the duration of treatment with Levodopa (main medication for PD treatment, Section 2.1.4) [31], although this association is not yet well understood. However, these episodes are more severe when the medication effect decreases - patient is in the OFF state (6-8h after the medication intake) [16].

Some patients develop tricks to overcome FOG episodes, which may include marching to command or step over objects [9]. It has been demonstrated that external rhythmic cueing through visual (e.g. floor projected color lines) or auditory (e.g. music, metronome ticking sounds) stimulation can help patients to overcome freezing attacks and resume normal gait [11, 12]. The auditory cueing, as it is easier to implement and has demonstrated good results, is the most implemented method.

Table 2.1: Description of the four main models that explain the origin of FOG. Reproduced from [41].

Models of FOG	Principle	Prediction of FOG-episodes
Threshold	Accumulation of motor deficits until threshold is reached and freeze occurs	Increase motor cycle frequency Decrease amplitude Increase coordination complexity
Interference	Competition for common central processing resources induces breakdown	Increase number concurrent tasks Increase difficulty level tasks Increase load of external input
Cognitive	Deterioration in processing of response conflict induces block	Increase incongruency level Increase response need Increase load on executive function
Decoupling	Decoupling between motor programs and motor response induces block	Increase strength startle stimuli Increase frequency startle stimuli Increase postural load or instability

2.1.2.2 Non-Motor Symptoms

Despite being considered a movement disorder, there are several non-motor symptoms associated with PD. Sometimes, these symptoms appear in early stages of disease, preceding motor symptoms, however they are not usually recognized as signs of PD. This is because its cause and relation with the PD is not yet well understood. Several studies are investigating these symptoms, looking for early diagnosis and better treatments [23].

The non-motor features comprise a heterogeneous set of symptoms. They include autonomic dysfunction, sleep and sensory disorders, cognitive impairments, fatigue, and others [9].

The **autonomic dysfunction** is related to the failure of the both central and peripheral autonomic nervous systems [23]. As such, several basic functions are compromised, leading to sweating abnormalities, gastrointestinal problems, bladder dysfunction, sexual problems and orthostatic hypotension - sudden fall in blood pressure when the patient stands up quickly - among others [9, 23]. These symptoms can appear in early or late stages of disease and significantly affect patients QOL.

Difficulties in sorting or planning tasks of the daily life are also reported by some patients. These issues are included in the **cognitive and neurobehavioral impairments** along with difficulties with thinking, word finding or judgment. These problems can evolve to dementia, increasing the burden of the disease to both patient and caregiver, being the main reason for nursing home placement [23].

Sensory abnormalities are common in patients with PD and are manifested through problems on olfactory function, loss or reduction in the sense of smell, that can be very useful in the early identification of the disease [42], and through a variety of pain syndromes that sometimes are so severe that can overshadow the remaining motor symptoms.

These patients also tend to suffer from **anxiety** and **fatigue**, which are closely related to **depression** and **sleep disorders**. Some sleep disturbances can appear as a side effect of the medication. Although, insomnia, sleep fragmentation, excessive daytime sleepiness and rapid eye move-

ment (REM) sleep - characterized by an increase in dream content - can be present throughout the various stages of the disease [43].

All of these symptoms really compromise the patient QOL, increasing the physical and emotional load of the disorder. In early stages, they may manifest in a mild way and thus, are usually overlooked and under appreciated as a PD symptom. With disease evolution, their severity increases, giving rise to serious complications. A better understanding of these symptoms and their relation with PD would allow an early and more efficient intervention.

Table 2.2: Summary of the most common symptoms of PD, reproduced from [9].

Motor Symptoms	Tremor, Bradykinesia, Rigidity, Postural Instability, Gait impairments (such as FOG), Slowness in ADL Hypomimia, Dysarthria, Dysphagia, Sialorrhea Decreased arm swing, Micrographia, Dystonia
Non-motor Symptoms	Autonomic dysfunction, Cognitive impairments, Depression, Fatigue, Apathy, Sensory impairments, Sleep Disorders

2.1.3 Clinical Assessment of Parkinson's Disease and Freezing of Gait

The gold standard for clinical assessment of PD is the application of clinical rating scales [4], namely the revised version of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS) [13]. It comprises a set of questions directed to the patient and caregiver and defines specific tasks that are visually examined and scored by the clinician. It prioritizes the motor aspects of PD, but also tries to capture the presence and severity of problems in the non-motor domain.

Another common scale is the Hoehn & Yahr (H&Y) scale, used to compare groups of patients and to provide a rough assessment of the disease stage. It is easy to apply and simply evaluates the severity of the disease based on bilateral motor involvement and gait and balance impairments [44]. Consequently, issues associated to other motor and non-motor symptoms are not completely captured. Originally, it had 5 stages (1-5), but was modified with the introduction of stage 1,5 and stage 2,5. The progression through these stages is highly correlated with motor decline and QOL [45]. Table 2.3 presents the various features of each stage.

Table 2.3: Stages of the Hoehn & Yahr scale updated version, reproduced from [45].

Hoehn & Yahr scale	
1.0	Unilateral involvement only
1.5	Unilateral and axial involvement
2	Bilateral involvement without impairment of balance
2.5	Mild bilateral disease with recovery on pull test
3	Mild to moderate bilateral disease; some postural instability; physically independent
4	Severe disability; still able to walk or stand unassisted
5	Wheelchair bound or bedridden unless aided

Regarding clinical evaluation of the presence of FOG, it can be performed through the application of specific FOG questionnaires [14] or direct observation. The former is similar to the ones described above, based on information gathered from the patient and caregiver about the severity and frequency of freezing episodes. However, all these questionnaires and rating scales can be unreliable, since they are based on subjective perspective of patients, caregivers and clinicians. The self-assessment is also inappropriate due to cognitive impairments or memory loss of the patients [15].

On the other hand, the direct observation consists in the performance of specific gait tests in order to elicit FOG. This is particularly challenging in the clinical and research environment and thus, some studies have investigated which are the most effective sets of tests to provoke FOG [46, 47, 48, 49]. The tasks include walking through a narrow space, walking with short steps rapidly, narrow turns and Timed Up & Go tests (the patient is timed while he/she rises from a chair, walks three meters away, turns and walks back to the chair, sitting down again [50]). In addition, Ziegler *et al.* [49] have created a simple protocol "Ziegler Protocol" with a defined set of sequential actions to provoke FOG in clinical settings. This protocol has been adopted by various authors [5, 8] and it is described in Figure 2.1. Tests can be accompanied by video recordings in order to perform posterior evaluations.

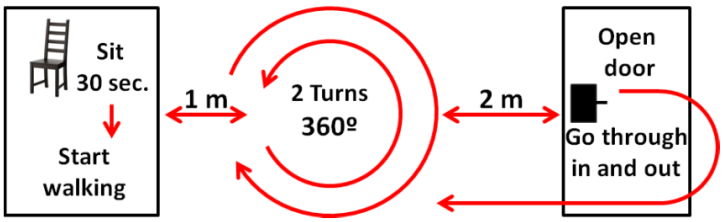

Figure 2.1: The Ziegler Protocol exercises: the patient is asked to sit down on a chair during 30s. Then, he/she is asked to stand up and walk to a floor mark, where he/she performs two 330º turns (clockwise and counter-clockwise). After that he/she walks to the door, opens it and passes through the door. He/she turns outside and comes back to the chair. (Reproduced from [8].)

This second approach has also some major drawbacks: the patient is not in his/her natural home environment and the test does not replicate real life settings. Moreover, the frequency and duration of the freezing episodes cannot be transposed to the daily life [18]. So, an in-home evaluation would be desirable, improving the reliability of the tests [15].

2.1.4 Treatment and Medication

Nowadays, there is no definitive cure for PD, only treatments that help patients to overcome symptoms. The oldest and most common pharmacological therapy for PD is Levodopa (Ldopa) [4, 9]. This drug is converted to dopamine, compensating the concentration lack of this substance in the brain. The introduction of Levodopa in PD treatment has significantly improved the QOL and decreased mortality [22]. Other dopamine agonists drugs have also been introduced.

However, the Ldopa therapy has several side effects, associated to motor complications, particularly dyskinesias - uncontrolled involuntary movements [4]. The induced dyskinesias are more likely to occur in younger patients [9] and so, in these cases, it is advantageous to delay the intake of medications such as Levodopa. In older patients, the risk of motor complications is smaller, when compared to the burden of the symptoms, and then the effect of medication is better [22].

The effect of medication on PD symptoms decreases over time leading to a ON-OFF phenomenon. When the patient is under medication effect, one could say that he/she is at ON state. On the other hand, he/she is at OFF state when the medication effect is gone and he/she is waiting to take another dose (6-8h after the last medication intake), giving rise to motor response fluctuations [16].

According to Giladi N. [35], most common FOG episodes are related to an OFF state, when the dopamine concentration is lower. However, there is no consensus regarding the effect of medication on this symptom. In [51], Giladi concluded that FOG is a heteregenous symptom that can be divided into dopaminergic sensitive, resistant or provoked by Levodopa. The patient treatment should be defined according to the type of FOG manifested by each patient.

Therefore, during the first years, while the medication produces significant benefit, PD patients are able to carry on with their ADL and live an almost normal life without additional assistance. Over time, the severity of the symptoms increases, the medication effectiveness is reduced and QOL is significantly affected. Physical activity, through rehabilitation exercises, mainly in the advanced stages, as well as psychological support, assume an important role during the disease [22].

2.2 Inertial Sensor Systems

In recent years, there has been an effort on developing new solutions for the monitoring of human motion during the performance of tasks of ADL, in unsupervised environments and in a continuous and non-invasive way. Most of these solutions have employed wearable devices, namely wearable inertial sensors. Usually, these sensors are based on Micro-Eletro-Mechanized-Systems (MEMS) and are combined together as Inertial Measurement Units (IMU) [52]. Several studies have applied this type of sensors to measure the motor performance of the PD patients, particularly to the assessment of freezing episodes, which proved to be crucial to improve the monitoring, assistance and evaluation of these patients [4, 6].

Due to the advances of the MEMS, these devices have become much smaller, compact and low cost. They are lightweight, non-invasive, easy and comfortable to be worn, enabling ambulatory monitoring [53] and improving user acceptance, in contrast to traditional solutions, such as optical systems and pressure platforms [4, 17]. They have low-power consumption and are able to send information to external devices, via wireless, allowing portability. Additionally, they are appropriate for real-time applications and can be easily integrated in alarm systems [53, 54], which can have a critical role in the management of PD symptoms, as stated before. Inertial sensors are

relatively accurate in measures, although they have associated drift errors, that can be corrected through the use of sensor fusion techniques [54].

Taking into account the features and advantages of inertial sensors, in this dissertation they are used to perform the gait monitoring of PD patients, aiming to detect FOG episodes and pre-FOG patterns. In the following section, a more detailed description of the Inertial Sensors Systems is presented.

2.2.1 Inertial Measurement Unit

IMUs have been widely used for monitoring human movements in different environments, from the clinical and laboratory environment to the real-life conditions. The human movement analysis has an important role in the evaluation of QOL and in the assessment and monitoring of activities and symptoms of patients with PD.

These devices usually include three types of sensors: accelerometer, gyroscope and magnetometer.

- **Accelerometers** are electromechanical devices that measure acceleration forces along one, two or three orthogonal axes. Usually, an IMU contains a 3 axis accelerometer, which gives the values of acceleration in m/s^2 along the X,Y,Z axes. This type of sensors is used to recognize motion. The readings of an accelerometer are subjected to errors, being the bias the main source of them. The bias is the offset of its output signal from true values [55].

- **Gyroscopes** provide values of angular velocities in rad/sec along the three axis, X,Y,Z, and could be used to get correct orientation of the device while in motion. Gyroscopes sense orientation through changes in the values of angular velocity. However, they have a tendency to drift over time, since they have no fixed frame of reference [56]. It is subjected to bias and numerical errors. The bias is the offset of the output from the true value and appears after the integration as an angular drift, increasing linearly over time. Calibration errors are also associated to the measures [55].

- **Magnetometers** provide values of the local magnetic field vector components, in micro Tesla. They can be used in combination with accelerometers to find the direction with respect to North when linear acceleration is zero. Magnetic interferences are the main source of errors in these devices [54, 55].

These measures are obtained along the axes of the coordinate system fixed with the device [54]. Thus, orientation changes of the device can influence the collected values. This can be particularly challenging during the motion and gait analysis using IMUs, since the relative orientation between sensors and subject body cannot be fixed and it may change with body movements [52]. In these cases, extracting orientation invariant motion features could be advantageous.

Additionally, for data analysis purposes, one can use raw data, obtained from each individual IMU sensor, or combine data of the different sensors, in order to obtain more information in a

more accurate way. This is achieved through the application of sensor fusion algorithms. In the case of IMUs, they are mainly used to optimize device orientation estimation. This information can have great interest in the calculation of orientation independent measures.

2.3 Use of Technology for Assessment of Freezing of Gait

As referred in Section 2.1.2, FOG is one of the most disabling symptoms of PD, characterized by a block of movement during gait initiation, turning, narrow quarters or before an obstacle [48]. It is a common cause of falls and significantly decreases patients QOL. The current assessment of this symptom, through the application of specific questionnaires, is not accurate and may not reflect its severity, in terms of frequency and duration of episodes [57]. Regarding these limitations, several studies have applied new technologies to perform an objective and reliable assessment of FOG. As stated by Rodriguez-Martín D. *et al.* in [18], on their related work revision, some of these studies include the use of pressure plataforms [58], optical systems and sensors to monitor signals such as electroencephalogram [59], eletromyography or skin conductance [60]. Besides the relevant information gathered from these methods, they do not allow its application outside clinical environments, making continuous monitoring impossible, and they have low acceptance among patients. Thereby, non-invasive wearable systems have gained an important role in monitoring freezing episodes, mainly the ones based on inertial systems [4].

The following sections present an overview of the major developments performed in the research field for the assessment of FOG. They focus on the studies that apply inertial systems, since they are used in the present work. The studies are divided in the ones that detect freezing events, when the FOG is happening, Section 2.3.1, and the ones that aim to predict when a FOG is about to happen, Section 2.3.2. Finally, to understand the commercial viability of this type of solution, an overview of the available systems in this field is presented, Section 2.3.3.

2.3.1 Detection of Freezing of Gait

Inertial sensors systems have been used in a large number of studies to evaluate motor symptoms of PD, namely FOG. They are used for the detection and analysis of these episodes, in order to characterize its severity and to enable the application of rhythmic auditory cueing, that acts in the shortening of FOG duration. For this purpose, a minimum latency is required. Studies differ in the type and number of sensors, their location in the patient body, as well as in the chosen methodology for data processing. Table 2.4 summarizes the main studies for FOG detection.

The first study aiming the detection of FOG episodes belongs to Han *et al.* [61], in 2003. They used 3-axis accelerometers attached on the ankles to analyze features of the freezing and normal gait.

In 2008, Moore *et al.* [62] performed a study with 11 patients with advanced PD. They tried to validate ambulatory monitoring of FOG using the frequency characteristics of vertical leg movement. Signals were acquired through accelerometers placed on the ankles of the patients. In this study, they introduced a new concept, the Freezing Index (FI), defined as the power in the *freeze*

Table 2.4: Overview of the main studies related to the detection of FOG

Author	Year	Sensors	Location	Patients	Methods	Contributions / Main Results
Han *et al.* [61]	2003	2 Accelerometers	Ankles	2	Statistical test	The frequency response of a FOG remains in the 6-8Hz band
Moore *et al.* [62]	2008	1 Accelerometers	Left Ankle	11	Threshold based	Introduction of new concepts: freeze band, locomotion band and FI. Application of a general (78.3% success) and a patient adjusted threshold (89.1% success).
Bächlin *et al.* [63]	2009	1 Accelerometers	Left Ankle	10	Threshold based	Introduction of the concept of PI and a second threshold. The first study to perform automatic online detection of FOG. Sensitivity: 73.1% Specifity: 81.6%
Mazilu *et al.* [64]	2012	3 Accelerometers	Lower back, Shank, Thigh	10	Machine learning	The first study applying machine learning techniques. Several classifiers and temporal - frequency features. Sensitivity: 98.35% Specificity: 99.72%
Moore *et al.* [57]	2013	7 IMUs	Lumbar back, Thighs, Shanks and Feet	25	Threshold based	Evaluation of different sensor locations, windows size and threshold values. Seven-sensor configuration - sensitivity: 84.3% specificity: 78.4%
Mazilu *et al.* [8]	2014	2 IMUs	Ankles	5	Machine learning through a decision tree classifier	Able to be used in unsupervised environments. User independent and real time methodology. Sensitivity: 97%; 27 false positives; 99 true positives.
Ahrichs *et al.* [65]	2015	1 Accelerometers	Waist	8	Machine learning through SVM	FFT, FI and time-frequency features. Application of a degree of confidence to the classifiers output. Sensitivity: 91.1% Specificity: 100%
Rodriguez-Martín *et al.* [18]	2017	1 Accelerometers	Waist	21	Machine learning through SVM	Generic and patient-dependent approaches were developed. 55 features explored. Better results with the SVM patient-dependent methodology. Sensitivity: 88.1% Specificity: 80.1%
Lorenzi *et al.* [66]	2017	1 IMU	Head	-	Machine learning through the use of ANN	Not satisfactory results for PD patients.
		2 IMU	Shins	16	Threshold based	Sensor fusion techniques. Quaternion based representation of the limb orientation and position. Robust for detection of FOG. Sensitivity: 94.5% Specificity: 96.7 %.

band (from 3 to 8 Hz) divided by the power in the *locomotor* band (from 0.5 to 3 Hz). During FOG, the power is concentrated in the first band, whereas voluntary activity exhibited significant power in the last one. Hence, a FOG event occurs when this ratio exceeds a *freezing* threshold. The value of the threshold could be global or patient adjusted, the last one resulting in a better performance. One year later, Bächlin *et al.* [63] performed for the first time an automatic online detection of FOG. Similarly to the previous study, they used accelerometers placed on the ankles to evaluate the frequency components of motion. Additionally, they tried to enhance the performance of the algorithm proposed by Moore *et al.* [62], by reducing the latency and decreasing the number of false positives when the patient is at rest. This was achieved through the introduction of a new parameter, Power Index (PI), that is the total energy between 0.5 and 8 Hz, indicating the amount of movement, and a second threshold, needed to avoid that standing parts are detected as FOG. Thus, a freezing event is detected when the quantity of movement PI is above the second defined threshold and the FI exceeds the *freezing* threshold. These modifications are needed for online purposes, enabling the application of auditory cueing, which started with a maximum delay of 2 seconds. In this study, the performance of the system in different locations (thigh and lower back) was also investigated, concluding that all of them could be used with almost no differences in the FOG detection performance.

The algorithm, presented by Moore, and enhanced later by Bächlin (MB) [63], has been employed in several studies, being the most used approach for FOG detection and for comparison purposes [18, 19, 57, 64, 67, 68].

In 2013, Moore [57], based in the previous approach, evaluated the effect of sensor locations and signal parameters, such as freeze threshold and sampling window size (for determining the ratio between the *freeze* and *locomotor* bands), on the performance of FOG detection. It was concluded that windows with a longer size result in better performances, although they also lead to a loss of sensitivity at episodes with short duration. Seven sensors attached to the lumbar back, thighs, shanks and feet were explored in a group of 25 patients. The configuration comprising seven sensors has proved to be the most accurate method, although it lacks in comfort and usability. On the other hand, single shank sensor results are very similar to the ones obtained by the seven-sensors approach, being much simpler to use.

Different methodologies based on machine learning classifiers have been developed, encouraged by the successful applications in the field of activity recognition [69]. FOG can be seen as an involuntary activity, characterized by typical motion patterns different from normal gait and thus these techniques can have here a critical role [64].

Mazilu *et al.*, in 2012 [64], were the first ones to apply machine learning techniques to the online detection of FOG, using wearable accelerometers and a smartphone. They also included auditory feedback every time a new event is detected, with a mean latency of 0.34 seconds. They explored several machine learning algorithms (Random Trees, Random Forests, decision trees, Naive Bayes, k-Nearest Neighbor, and others) with different temporal-frequency features, sensors location and windows size to optimize FOG detection accuracy and latency. Results obtained demonstrate the potential of machine learning approaches.

Although all these studies have achieved good performance in the detection of freezing episodes, most of them were tested under controlled conditions. This suggests that their performance could decrease if they were applied during unsupervised daily activities. This being so, more recent studies have tried to develop methodologies that can be transferable to outside the clinical setting.

In 2014, Mazilu's group [8] followed the findings and methods presented in [64] and developed a wearable assistant that patients can use in daily life, in unsupervised environments, without additional assistance - GaitAssist. System has undergone a design phase where the authors received feedback from PD patients, clinicians and engineers, to define the technical and usability requirements of the system. It consists of two IMUs attached on the ankles and a smartphone that collects and processes sensor data. During data processing, several features, such as FFT-based features, mean, standard deviation and movement amplitude were extracted from acceleration magnitude signal. Features were fed into a two class decision tree classifier, with FOG and Normal Gait classes. This is a user independent model that can be executed in real-time, making possible auditory cueing that helps the patient to overcome freezing episodes. The system was tested in 5 different patients and was able to identify 99 of 102 FOG events. Auditory cueing started with a maximum latency of 0.5 seconds.

In the study performed by Ahrichs *et al.* [65], Support Vector Machines (SVM) were used for detecting FOG from data recorded while patients performed several scripted activities in their home. The proposed method is user independent and based on acceleration measurements obtained from a waist-worn device. Two feature sets are evaluated: (i) a reduced set with only the Fast Fourier transform (FFT) and (ii) a full feature set comprising FI and time-frequency features. These are extracted from equally sized windows and fed the SVM classifier. Outputs are aggregated for the calculation of a degree of confidence and if it exceeds a threshold, the aggregated time frame is considered a FOG. A second threshold, in order to set a maximum and minimun value of confidence for freezing episodes, was also explored. Despite high values of accuracy obtained, this evaluation could last almost 60 seconds, which makes impossible the real time detection and cueing application.

Rodriguez-Martín *et al.* have also explored a machine learning approach based on SVM for online detection of FOG in patient's own home [18]. Data were obtained by a single accelerometer placed on the left side of the waist of 21 PD patients. A generic and a patient-dependent approach were developed and the performance was compared with the methodology purposed by Moore and Bächlin [63]. Results obtained show that personalized models, for both SVM and MB method outperform generic models. The performance of SVM approach is also better. For the input on the SVM, 55 features were extracted from recorded data. In another study [15], authors were able to decrease the number of features from 55 to 28, reducing computational burden of the algorithm and improving real time detection. In [17], they present the design of a custom IMU optimized for data acquisition and storage and for real-time execution of the implemented machine learning algorithm, the SVM. They tested the new device in 12 patients and achieved good results.

A wearable wireless sensing system for monitoring motion disorders was also proposed by Lorenzi and Irrera in recent works [20, 66]. They used wearable inertial sensors to identify motion

disorders, not limited to PD, in two different configurations: a single sensor on a headset and two sensors on the shins. The first configuration uses Artificial Neural Networks (ANN) to recognize the patterns of the gait. Despite being effective when used in patients with temporary traumas at the limbs, this methodology has shown to be unsuitable for patients with PD, because they suffer from a variety of postural issues that overlap the signals that identify the patterns of FOG. On the other hand, the second configuration, based on time domain analysis of sensor signals, after applying sensor fusion techniques on raw data, is able to sense all the movements in the lower limb. It has shown to be robust and reliable for PD patients, distinguishing between freezing and normal gait.

2.3.2 Prediction of Freezing of Gait

The studies presented above detected FOG after or while it is happening. However, there are some evidences suggesting that prior to freezing, characteristics of gait can deteriorate until a breakdown of locomotion is achieved [37, 41, 70].

Based on this hypothesis, a few researches have been conducted aiming FOG prediction instead of FOG detection, and are summarized in Table 2.5. It could be a challenging approach, that can have great benefits. Instead of acting on helping patients to resume walking after a freezing period, one can apply a preventive cueing reducing the probability of a FOG to occur.

Some attempts have been performed regarding early detection of FOG, like the study of Handojoseno *et al.* [59] where they analyzed brain activity during the onset of FOG. In 2013, Mazilu and his group [19] were the first ones to handle FOG prediction using motion data acquired from 3-dimensional accelerometers placed on the ankles. They explored unsupervised feature learning based on principal component analysis, both for FOG detection and prediction, and compared results to other feature learning approaches based on standard frequency, time-domain and statistical features. A Decision Tree Classifier was adopted due to its low computational cost. For FOG prediction a three-class problem was formulated (FOG, pre-FOG and normal gait), obtaining highly patient-dependent results.

Two years later, Ferster *et al.* [71] used data from accelerometers and gyroscopes placed on the ankles of 5 patients and characterized gait parameters during the transition from normal walking to freezing episodes. Two types of features were explored: (i) gait based features and (ii) frequency based features. For the first type, gait cycles from IMU raw data were detected and then features as stride duration and stride length, among others, were extracted. For the last one, raw IMU data were separated in overlapping windows for feature extraction. Their results suggest that there are variations of the features before FOG, the more relevant ones in the stride duration, stride length, dominant frequency and the inverse of the dominant frequency slope. Hence, it was concluded that this type of parameters might be used to predict FOG.

More recently, Palmerini *et al.* [5] also tried to identify pre-FOG phase through the study of gait before FOG. They used accelerometers and gyroscopes attached on the shins and lower back to identify and quantify characteristics of gait 2 seconds before block of movement. Eight features were extracted from the acquired signals and then subjected to statistical analysis in order to find

Table 2.5: Overview of the main studies related to the prediction of FOG

Author	Year	Sensors	Location	Patients	Methods	Contributions / Innovative Aspects
Mazilu *et al.* [19]	2013	1 Accelerometer	Ankle	8	Machine Learning through Decision Tree Classifier	The first ones to handle with FOG prediction. Unsupervised feature learning were explored. F1-measure of 56%
Ferster *et al.* [71]	2015	2 IMUs	Ankles	5	Analysis of features variation	Characterization of particular gait parameters before freezing.
Palmerini *et al.* [5]	2017	2 IMUs	Shins	11	Machine Learning through Linear Discriminant Analysis	Statistical Analysis applied on the extracted features. Sensitivity: 83% Specificity:67%

the ones that better discriminate between gait and pre-FOG. These were used to train a linear discriminant analysis classifier. Results obtained are patient dependent, but demonstrate that there is a potential on creating automatic algorithms able to identify and quantify periods before FOG, making possible its prediction.

2.3.3 Commercially Available Systems for Assessment of Parkinson's Disease

Thanks to developments in wearable technologies, there are already a few commercially available systems that aim to provide an objective and continuous assessment of some of the motor symptoms of PD. As far as it is known, none of these systems is capable of continuously monitoring FOG, yet they represent the ultimate goal of most of the presented studies: the possibility of creating an integrated system that can be used by PD patients during their ADL in home environment, helping them and their doctors to improve the management of the symptoms. Next, three of those systems will be presented.

The first one is the Parkinson's KinetiGraph™ [72], developed by Global Kinetics Corporation. The system provides an assessment of tremor, bradykinesia, dyskinesia, daytime somnolence and an indication of propensity to impulsive behaviours. It is also able to manage the medication intakes by reminding the patient at his/her medications times and asking him/her for a confirmation when a dose is taken. The device is similar to a watch, being a wrist-worn device with a 3-axis accelerometer able to record 11-bit digital measurement of acceleration with a range of $\pm$ 4g and with 50 Hz as sampling frequency [73]. Data may be recorded during 6 to 10 days and at the end of this period it is uploaded to be processed by the company. Then, the patient, or his/her clinician, should receive a detailed report regarding the patient movements. Figure 2.2 describes the main features of the system. Regarding data processing, frequencies between 0.2 and 0.4 Hz are analyzed through spectral power, peak acceleration and amount of time with no movement [73].

The Great Lakes Neurotechnologies company have also created two solutions: Kinesia ONE™ and Kinesia-360™. The former consists on a single inertial sensor placed on the finger (see Figure 2.3), aiming the assessment of PD symptoms, such as tremor, bradykinesia and dyskinesia, during specific defined tasks. During a day, the patient interacts with the system through an iPad, and follows a set of exercises, that will allow symptoms evaluation. Data are sent to a cloud storage able to be analyzed by the clinician [74].

On the other hand, Kinesia-360™ comprises two wearable inertial sensors, which include a 3-axis accelerometer and a gyroscope, placed on the wrist and on the ankle. It measures tremor, dyskinesia and mobility and also provides an electronic diary to record patient reported outcomes and medication. The interface between the user and the system is achieved through a smartphone application, Figure 2.3. Data are sent to an external server and data processing is performed based on temporal and frequency features [75].

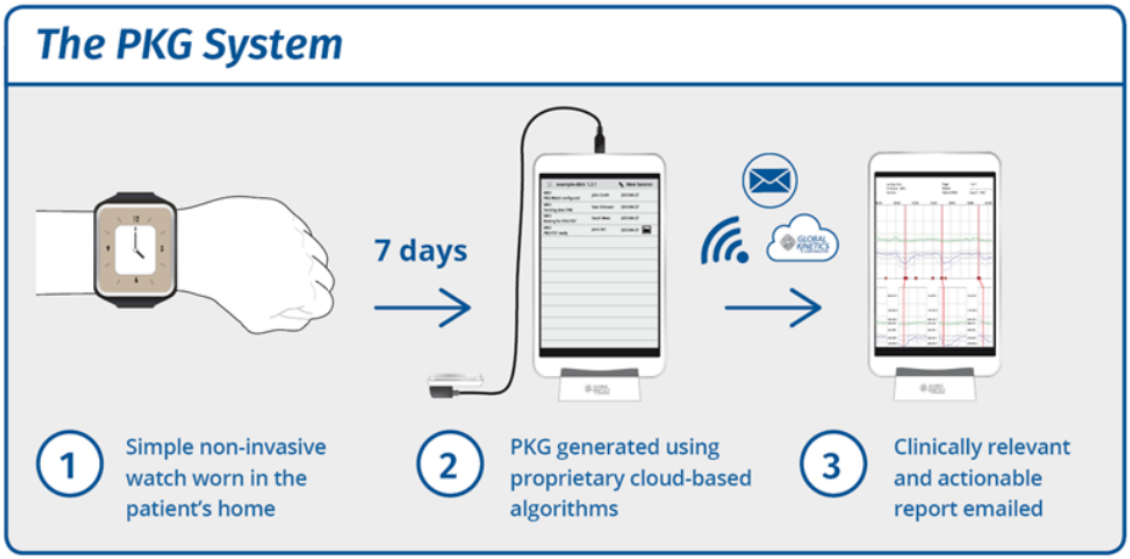

Figure 2.2: Parkinson's KinetiGraph™ system components [72].

Figure 2.3: Kinesia ONE™ (left) and Kinesia-360™ (right) systems [74, 75].

2.4 Gait Analysis

As mentioned before, inertial sensors systems have been widely used for monitoring human activity. The monitoring of FOG in PD patients is one of many applications. Data from accelerometers, sometimes together with gyroscope data, are frequently used to perform gait analysis. In this way, the understanding of the main gait features and of the corresponding acquired signal is important to be able to interpret it and explore the methodologies of analysis. Hence, in Section 2.4.1 a brief description of gait phases is presented and in Section 2.4.2 an analysis of the typical signals is performed.

2.4.1 Gait Phases

Bipedal walking, or simply gait, is an important basic feature of the human condition. It is a complex activity, characterized by smooth, regular and repetitive movements [76]. To generate walking, one needs to be able to perform the initiation and termination of locomotive movements, to continue the movement progressing toward a destination, while maintaining the equilibrium and adapting to any changes in the environment [77]. The study and characterization of human locomotion - gait analysis - mainly involve the quantification and description of the lower limb movements and it is applied in a variety of fields, such as biometrics, sports, orthopedics, rehabilitation, diagnosis and ambulatory monitoring [78]. The monitoring of gait of PD patients for the assessment and evaluation of freezing events is an important application of gait analysis.

The simplest form of gait is defined as gait cycle and comprises a single sequence of events performed by one limb, from the moment a foot touches to the ground to when that same foot touches again [77, 79]. The gait cycle can be divided in different phases and events, Figure 2.4. The two main phases are [79]:

- the stance phase - involves about 60% of the gait cycle and corresponds to the period when the foot is in contact with the ground. It begins with the initial contact (heel contact) and ends with the toe off of the same foot.

- the swing phase - accounts for 40% of the gait cycle and is the period when the foot is not on the ground. It begins when the foot is lifted and ends when the same foot contacts the ground.

The gait analysis is performed through the identification of gait events and determination of gait parameters, which can either vary among individuals or for the same subject, depending on the physical condition. Some parameters are the distance between the consecutive contacts of the same foot, called as stride length, the distance between heel contact of one foot and the contact of the other - step length - and the cadence, defined as the number of steps taken per unit of time.

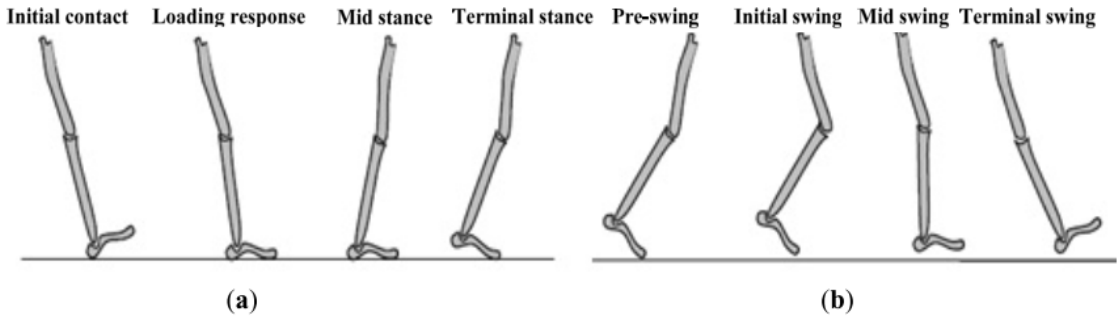

Figure 2.4: Gait phases in a normal gait cycle. (a) Gait sub-phases of the stance phase. (b) Gait sub-phases of the swing phase. Reproduced from [78].

2.4.2 Inertial Sensors Systems and Gait

Since the gait is an activity mainly performed by the lower limbs, the gait monitoring is usually performed with sensors placed at the ankles, legs and trunk. The acquired signals are different in each location and from them the identification of the main phases and events of gait is possible.

Figure 2.5 presents the relationship between the acceleration modulus (the modulus computed with the three axis x, y z) acquired with a sensor placed on an ankle and the gait events that belong to a gait cycle.

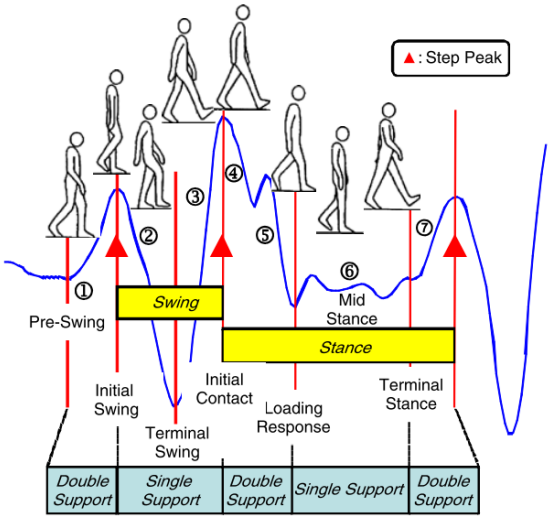

Figure 2.5: The relationship between the ankle acceleration during gait cycles and the temporal gait parameters for each gait cycle. Reproduced from [80].

When the patient initiates foot contact with the ground there is a big peak of acceleration, which magnitude is related with the strength of the heel contact. Usually, it is the bigger peak in the acceleration data, represented by number 4, in Figure 2.5. While the two feet are in the ground - double support - one can see a deceleration phase, identified with number 5, followed

by an almost constant acceleration period, number 6, corresponding to the moment when only the considered foot is in the ground - single support. A second peak in the signal is observed when the foot is off the ground and represents the end of the stance phase. It is followed by an acceleration valley, numbers 2 and 3, corresponding to the swing phase, when the considered foot is not in the ground.

2.5 Closing Remarks

Through the analysis of the fundamentals associated with PD it was concluded that freezing episodes significantly affect the normal living of the patients, having a negative impact in their QOL. Since there is no cure, the efforts are concentrated in developing new strategies to help patients and caregivers to better deal with PD and overcome FOG episodes. Although current assessment of FOG, performed through clinical rating scales, has an important role in the evaluation of the symptom, in most cases it is shown to be not reliable and ineffective, giving space for a subjective and inaccurate mapping of patient status. Hence, the need is raised for new objective strategies to manage freezing symptoms. These strategies require the possibility of monitoring the frequency and duration of FOGs, in patients home environment, with the important addition of being able to help patients handling with periods of block. Help may also be provided through rhythmic auditory stimulation that is activated during an event of freezing or even right before its occurrence.

Developments in wearable technologies allowed the appearance of new solutions and advantages of the existing Inertial Measurement Units making them the most promising technology in the field of activity recognition and thus in the monitoring of FOG.

The review of the studies performed in this area reveals promising works, some of them with a high probability to leave the research field and start to be used in real life conditions. Regarding the prediction of FOG, preliminary results have associated limitations. However, they provide evidences that freezing events can be predicted through the application of automatic algorithms, opening the door to further developments.

Chapter 3

Methodology for Detecting Freezing of Gait episodes

This chapter presents the development steps of the methodology for FOG detection proposed in this dissertation, as well as the dataset that was used.

Since a publicly available dataset was the only one used and tested in this dissertation, a good understanding of the acquisition conditions and protocol is fundamental. Thus, Section 3.1 presents a detailed overview of it. Additionally, a preliminary analysis of the dataset signals is also performed and presented in Section 3.1.1.

Then, based on the literature review presented before and taking into account the preliminary analysis, a set of steps aiming FOG detection were designed and are presented in the remaining sections. The implementation pipeline follows the traditional main stages to solve this kind of problem: signal pre-processing (Section 3.2), signal segmentation (Section 3.3), feature extraction (Section 3.4) and classification (Section 3.5). In some of these steps, approaches proposed in the literature were adapted or even adopted and, in others, different approaches were created. At the same time that the steps are presented, associated theoretical concepts, fundamental for the understanding of the proposed methodology, are also explained. The algorithm is implemented in Python language [81] using the PyCharm software [82]. Throughout the sections, the main Python modules needed for the implementation are identified.

3.1 DAPHNet dataset

In order to study the FOG phenomenon, data from a public dataset - DAPHNet dataset [16] - were used in this dissertation. It is important to notice that a data acquisition protocol was also designed within the scope of this project, aiming a new dataset construction. However, to perform this kind of study, an ethical approval is necessary. Hence, several documents had to be submitted to be approved by the Ethics Committee of the "Hospital S.João". Because the approval process took a long time, it was only conceded two weeks before the conclusion of this report. The documents

concerning the identification of the study and the final approval are presented in Appendices A.1 and A.2.

Consequently, the DAPHNet dataset was the only one used and tested in this dissertation. It was acquired with the aim of evaluating methods for the recognition of FOG periods from acceleration data collected from the lower limbs and has already been used in many studies concerning the FOG detection [19, 64, 83]. To build the dataset, 10 PD patients participated in the study. The patients (7 males and 3 females) had an average age of $66,4 \pm 4,8$ years and the average of disease duration was $13,7 \pm 9,67$ years. The Hoehn & Yahr score was in average $2,6 \pm 0,65$. All of the patients performed the tests in the OFF state, except for two patients that had an history of frequent FOG events during the ON state. This information is reported in Figure 3.1 for each patient.

Subject ID	Gender	Age [years]	Disease duration [years]	H&Y in ON	Tested in
01	M	66	16	3	OFF
02	M	67	7	2	ON
03	M	59	30	2.5	OFF
04	M	62	3	3	OFF
05	M	75	6	2	OFF
06	F	63	22	2	OFF
07	M	66	2	2.5	OFF
08	F	68	18	4	ON
09	M	73	9	2	OFF
10	F	65	24	3	OFF
Mean ± STD		66.4 ± 4.8	13.7 ± 9.67	2.6 ± 0.65	

Figure 3.1: Personal and disease information of the patients from the DAPHNet dataset, reproduced from [16].

During the study, three sensors were placed in the lower limb of each patient, namely at the ankle, at the thigh and at the trunk, in order to measure the acceleration, while performing a set of walking tasks that intend to represent different ADL. The sensors recorded at $64Hz$. It should be noticed that there is no information about the placement side (left or right) of the sensors or if care was taken to place the sensors always in the same side and using the same orientation. The walking tasks were performed in a lab and were defined aiming the trigger of many freezing episodes. The session was about 10-15 minutes and included the following phases, Figure 3.2:

- walking back and forth in a straight line, including 180 degrees turns,
- random walking, including initiated stops and 360 degrees turns,
- walking simulating ADL, such as entering and leaving rooms, getting something to drink and returning with a cup of water (multitasking).

The sessions were recorded and then FOG episodes were identified in a post-hoc analysis, to determine the exact start and end times and durations. The labeling was performed considering that the beginning of a FOG corresponds to the moment when the gait pattern starts to be arrested. The end of a FOG was defined as the point in time at which the pattern was resumed.

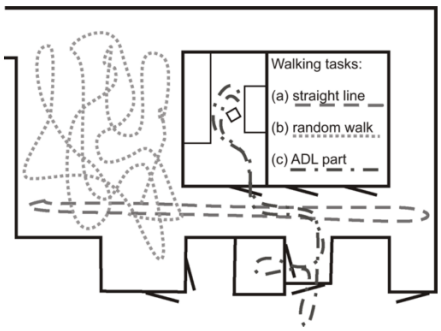

Figure 3.2: Sketch of the path taken by the subjects during the study, from [16].

In total, in the referred study, 237 FOG events were identified ranging from 0 - 66 per patient with a mean of 23,7. Patient 4 and 10 did not have any freezing episode. The authors reported that patients expressed different walking styles. According to them, Patient 1 suffers from foot drop while walking, which affects the patient ability to raise their feet at the ankle. Thus, the movement is characterized by intense stepping movements along the vertical axis. On the other hand, Patient 8 had the slowest and limited mobility, being the patient most affected by PD, Figure 3.1.

The dataset comprises a set of .txt data files, each one containing the following attributes:

- Time of sample in milliseconds;
- Ankle (shank) acceleration - horizontal forward acceleration [mili-gravitational acceleration - mg];
- Ankle (shank) acceleration - vertical [mg];
- Ankle (shank) acceleration - horizontal lateral [mg];
- Upper leg (thigh) acceleration - horizontal forward acceleration [mg];
- Upper leg (thigh) acceleration - vertical [mg];
- Upper leg (thigh) acceleration - horizontal lateral [mg];
- Trunk acceleration - horizontal forward acceleration [mg];
- Trunk acceleration - vertical [mg];
- Trunk acceleration - horizontal lateral [mg];
- Annotation [0, 1, or 2] : 0 - not part of the experiment (when the sensors are installed but the patient is not performing activities of the protocol). 1 - experiment, no freeze (can be any of stand, walk, turn). 2 -freeze.

The first steps performed on the available data were the acceleration units conversion to the international system units (meters per square second - m/s^2) and the removal of the samples that

were not part of the experiment (labeled with 0). The remaining labels were also changed to have no freezing samples labeled as 0 and freezing samples labeled as 1. The data concerning the no freezer patients (4 and 10) were not considered.

In this dissertation, the main focus was the data acquired from the ankle (shank) and thus all the presented acceleration signals of the following sections were recorded from sensors placed there. The presented results in Chapter 5 also concern the ankle acceleration signals. The first reason for this choice is that the ankle is the location closest to the feet, where the patient more strongly feels the freezing event (they describe it as their feet were glued to the ground). Additionally, the literature reports the ankle as a good location for sensors placement, leading to good results [57, 71].

3.1.1 Preliminary analysis of the signals: Gait vs FOG

This subsection performs a preliminary analysis of the data from DAPHNet dataset, by visualizing the acquired signals from the eight freezing patients. This kind of analysis is extremely important in the definition of the methodology for FOG detection. The main goals are:

- to inspect if there are different patterns of gait and FOG among the patients;
- to understand how a freezing period is characterized in terms of signal shape;
- to understand the main differences between a normal gait period (no freezing events) and a freezing period;

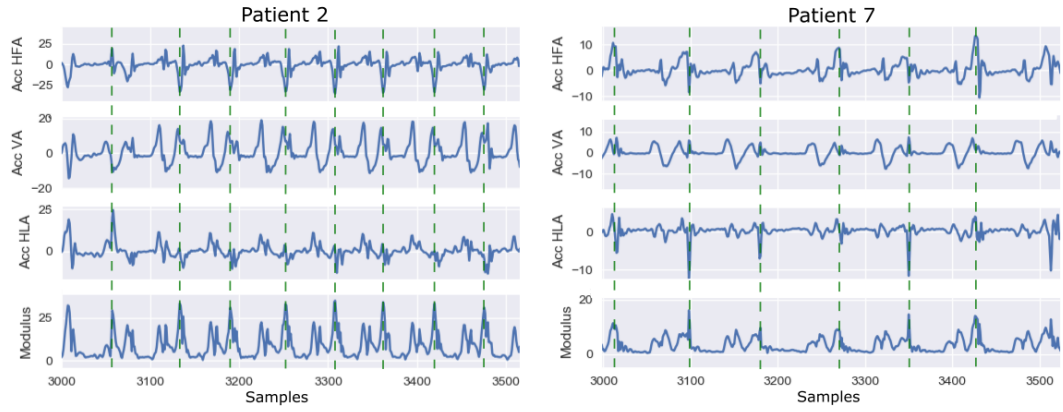

Figure 3.3: Comparison of gait patterns of two different patients. The acceleration values are in m/s^2.

In Figure 3.3, ankle acceleration signals from normal gait periods from two different patients are presented. The signals have patterns similar to the ones found in literature and thus, the main features of normal gait, such as the beginning of the stance phase, characterized by the initial contact, are easily identified - dashed green lines in Figure 3.3 . Nonetheless, the set of gait events

manifests differently in the patients. That is, each one has a different pattern of gait. The values of the acceleration at the initial contact and at the end of the stance phase vary among subjects as well as the duration of a complete gait cycle - cadence. In Figure 3.3, Patient 2 exhibits approximately 8 gait cycles during almost 8 seconds of signal recording, while Patient 7 exhibits 6 gait cycles in the same period of time. This could have a great importance in the following stages of signal segmentation. These variations are due to the natural differences in walking between different people, but, in this case, one can extract that age and stage of PD are the main contributors to the differences. Despite these variations, the signal of acceleration modulus is the one that presents more regular patterns and that allows an easier identification of the gait cycles, through the higher peak detection, corresponding to the initial contact moment. Thereafter, this signal will have an important role in the following stages of signal processing.

As mentioned before, FOG is characterized by a block of movement, a period when the patient attempts to progress the gait, unsuccessfully. Thus, the normal gait cycles are expected to appear modified or even unidentifiable. Figure 3.4 shows examples of normal gait data and complete FOG episodes from two patients of the DAPHNet dataset.

Observing the figures, the first conclusion to be drawn is that there is no specific pattern for the freezing events, not even for events of the same patient. Then, one can say that there is intra and inter patient variability. The episodes may have different durations and may manifest themselves in periods of low acceleration and near absence of movement and in periods with higher values of acceleration and abrupt and irregular transitions. The latter can correspond to repetitive movements, similar to tremors, performed during the attempt of resume walking. Sometimes, the two manifestations are visible throughout the same episode. It is important to notice that, the DAPHNet dataset gives no information about the context of each FOG, being impossible to know the conditions that lead to each episode, namely, if it happened during a turn, during a walking through a narrow space or during free walking, for example. This would be a valuable information since it would categorize the events and maybe different manifestations of FOG would be associated with different contexts.

3.2 Signal Pre-processing

The data from DAPHNet dataset consist in raw acceleration data, that need a first step of signal conditioning to prepare them for further processing.

The pre-processing stage has two main steps: signal detrending and signal filtering. The former was performed by removing the mean from the signal, which is an useful step for the further signal analysis in the frequency domain (for Fast Fourier Transform calculation). The signal filtering intends to remove high-frequency noise from the raw signal by implementing a second-order low-pass Butterworth filter, as proposed in the work of Martín [18]. The creation of the filter is done by executing the `butter(N, Wn, 'lowpass')` function available in Python, which designs a N^{th} order low pass filter, with a cutoff frequency Wn, returning the filter coefficients. Then `filtfilt` function applies the filter twice, forward and backward, so that the phase response of

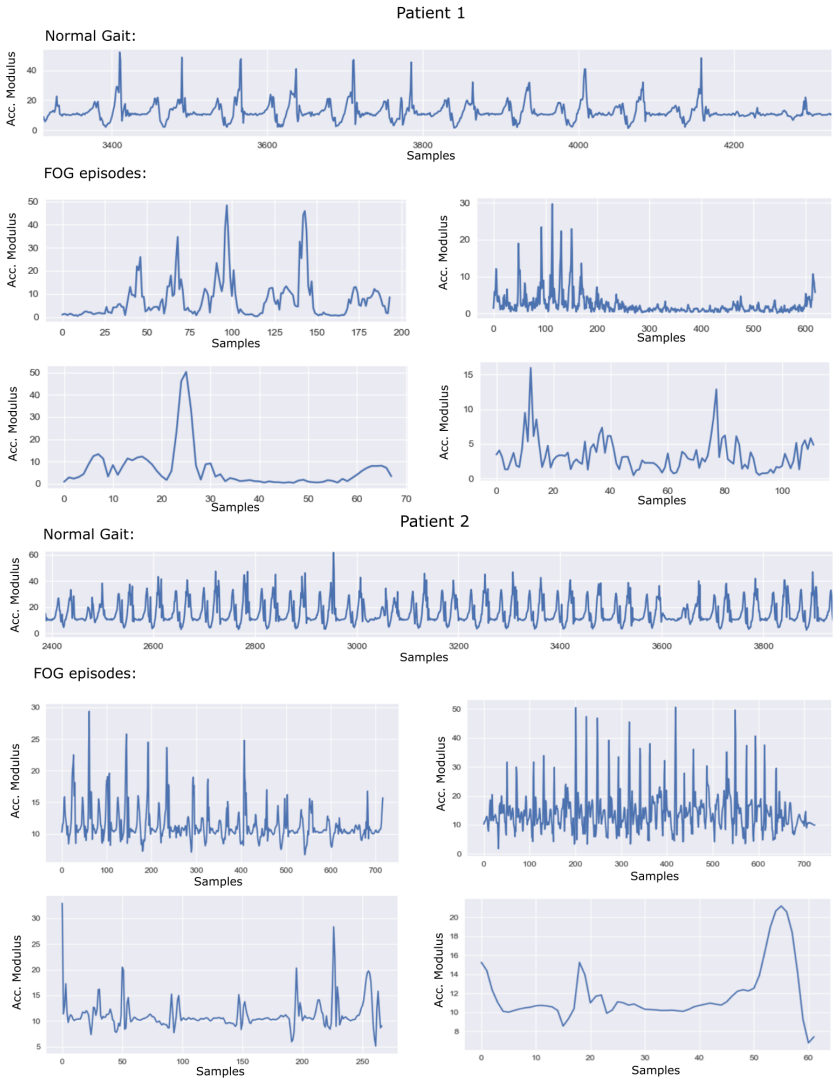

Figure 3.4: Examples of normal gait and FOG periods, from Patient 1 and Patient 2. There is a high variability in the manifestation of FOG either for the same subject or among individuals. The acceleration values are in m/s^2.

filter is linearized [84]. Both functions are available through the Python signal processing toolbox `scipy.signal`. Here, a cut-off frequency of 20Hz was defined experimentally and based on data from literature, in a way that the high frequency noise, not related to human movement, could be removed, without loosing relevant signal information from gait and FOG periods.

3.3 Signal Segmentation

The Signal Segmentation step intends to divide the signal into smaller portions - data windows. Most of the works analyzed in the literature review divided the signal into fixed sized windows, sometimes with no agreement with the best temporal window to be used. For example, in [57] the authors recommended to work with windows of over 2,5 seconds; the authors of the DAPH-Net dataset [16] worked with windows of 4 seconds and in [19] Mazilu *et al.* used windows of 1 second. After observing that the ideal window's size could depend on the concerned patient (since different patients have different values of gait cadence) and having in mind the features that will be extracted, as detailed in Section 3.4, a patient dependent segmentation is proposed in this dissertation. Hence, a patient adjusted temporal window is created, based on the size of the patient gait cycle.

Consequently, an algorithm to detect gait cycles based on the acceleration data becomes necessary. As analyzed in Section 3.1.1 Figure 3.3, the acceleration modulus signal, recorded from the ankle, gives a very simple pattern of gait cycles, which is easy to detect through the higher peak of acceleration identification, representing the heel contact moment. In order to identify the acceleration peaks, the `detect_peaks` function, available in [85], was used. It detect peaks in data based on their amplitude and distances. Thus, to detect the desired peaks the following parameters were set:

- minimum peak height *mph* - defined as the sum of the signal mean and the signal standard deviation: $mph = mean_{signal} + std_{signal}$
- minimum peak distance *mpd* - set as 40 samples, chosen experimentally.

The function returns a set of peaks, representing the instants of initial contacts and limiting portions of the signal - windows. However, the signal is not just a set of normal gait cycles. There are moments of absence of movement - when the patient is voluntarily standing - and, as seen before, moments of FOG, where the gait cycles are modified or unidentifiable. Figure 3.5, illustrates the result of the peak detector algorithm and the consequent division of the signal. As one can see, the created windows will not always represent individual gait cycles, making it necessary to perform a subsequent division of those sequences. This new division is done based on the mean size of the gait cycle of the patient being tested. So, if a previous created window is bigger than two times the mean size of a gait cycle, it is divided again in new windows with the mean gait cycle size.

Having the signal divided into data windows, their identification as a Gait (label 0) or a FOG (label 1) window is necessary for further classification steps. The identification is done based on the label of the samples belonging to each window: a window is labeled as a Gait window, if most of the samples are labeled as Gait. Otherwise, it is labeled as a FOG window.

This signal segmentation approach presumes the existence of a calibration step, for each patient being tested, to identify the mean size of a gait cycle and the signal mean and standard

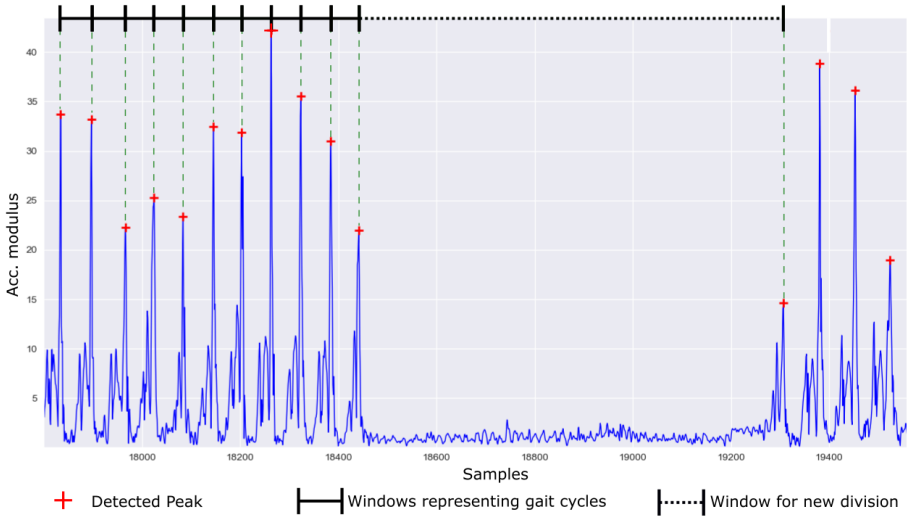

Figure 3.5: Illustration of signal segmentation process based on gait cycles identification, using a peak detection approach. The graph represents approximately 28 seconds of acceleration data recorded from patient number 1. The final windows representing individual gait cycles are identified as well as an example of a window that needs a further division, based on the mean gait cycle size. The acceleration values are in m/s^2.

deviation, needed to adjust the peak detector function parameters. Consequently, for each patient of the DAPHNet dataset the first ten consecutive cycles of normal gait were analyzed to perform the calibration. In a context of new data acquisitions, an initial moment of normal walking should exist, outside the FOG stimulating environment, to record calibration signals.

3.4 Feature Extraction

Concerning the Feature Extraction step, usually a large set of features, such as time and frequency domain features often used in activity recognition problems [19, 86], is extracted. However, for some of them, the reason why they are used and their relevance for FOG events detection is not clear. Hence, at this point, one of the goals was to explore more intuitive and meaningful features for the FOG detection problem.

3.4.1 Similarity Measures

Based on the signals variability during gait and FOG periods, features based on similarity measures are explored. Similarity measures on time series data are more difficult to compute, since the order of the elements needs to be considered [87].

In time series classification applications, the similarity measures are computed along with template matching approaches: a pair of time sequences are compared in order to identify parts of one sequence X, with size s, that match a predefined one, template Y, with size t, Equation 3.1.

$$X = (t_1, x_1), (t_2, x_2), ..., (t_s, x_s), Y = (t_1, y_1), (t_2, y_2), ..., (t_t, y_t) \tag{3.1}$$

In the context of FOG detection, the basic idea is to create a template of a gait cycle for each patient, based on the signal acquired during the calibration phase. Then, the created template will be compared with every signal sequence obtained in the signal segmentation step. Each comparison provides a real number that quantifies the similarity through a distance value. The smaller the distance is, the more similar the template and the sequence are [88]. Since a freezing event is a perturbation in normal gait, the comparison between the gait template and a FOG window is expected to lead to higher values of distance, being a potential feature to discriminate between Gait and FOG.

In this work, the Dynamic Time Warping (DTW) [89] was used as similarity measure. The following subsections briefly describe the fundamentals behind DTW computation (Section 3.4.1.1), explain the fundamentals of template creation (Section 3.4.1.2) and detail how it was implemented in the proposed FOG detection algorithm (Section 3.4.1.3).

3.4.1.1 Dynamic Time Warping

DTW is a shape-based similarity measure between two temporal sequences used in pattern recognition, time series classification and others [90]. The DTW finds the optimal alignment between two time series sequences, which may vary in time and speed, and captures flexible similarities by aligning the coordinates inside both sequences [87]. Figure 3.6 illustrates the application of the DTW in two time sequences, in contrast with the Euclidean distance calculation. The Euclidean distance is the simplest distance between two time series data, but its results are not useful since it is not able to identify identical sequences that are slightly shifted along the time axis. The DTW overcomes this limitation, ignoring global and local shifts in time dimension [91].

DTW uses a dynamic programming technique to find the optimal warping path between the sequence X and Y, Equation 3.1. In order to compute the distance, it first creates a $s \times t$ matrix - distance or cost matrix - where each element $D(i, j)$ is a cumulative distance of a minimum of three surrounding neighbors. Each cell of the matrix is computed as follows:

$$D(i, j) = Dist(i, j) + min[D(i-1, j), D(i, j-1), D(i-1, j-1)] \tag{3.2}$$

Once the entire matrix is filled, the DTW distance is calculated from the last element $D(s,t)$ to the first, $D(1,1)$. A greedy search is performed evaluating each cell to the left, down and diagonally to the bottom-left. Whichever of these three adjacent cells has the smallest values is added to the warp path. The search continues when $D(1,1)$ is reached [91, 92]. A single point of a time series can map to multiple points in the other time series, consequently, the two time series being compared could have different sizes.

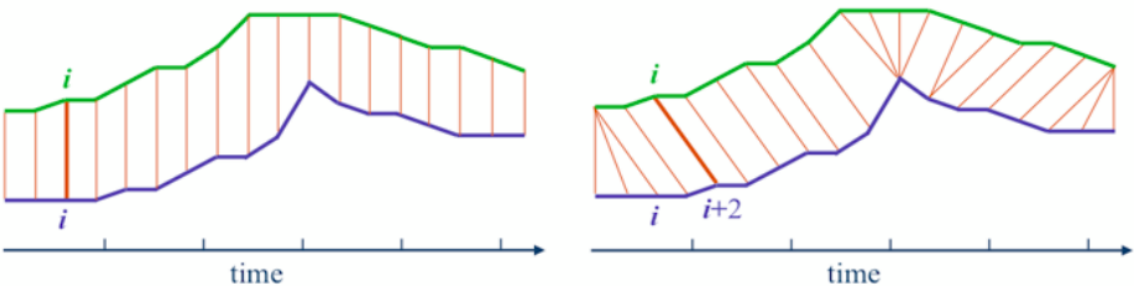

Figure 3.6: Example of similarity measures between two time series. At the left, the Euclidean distance, which aligns the i-th point of one time series with the i-th point of the other. It will produce a poor similarity measure. At the right, the DTW computation aligns points even if they are out of phase in time axis. Reproduced from [90].

A well know drawback of the DTW is the quadratic computational complexity [90]. The time and space complexity of the DTW is $O(s*t)$, since each cell in the $s \times t$ cost matrix is filled once in constant time. In this way, there are many attempts in literature to reduce time and space complexity, that falls into three categories [90, 91]:

- Constraints - It reduces the quantity of cells that are evaluated in the cost matrix;
- Data reduction - The dataset size is reduced before DTW computation;
- Indexing - Use lower bounding functions to decrease the DTW calculation in Data Mining applications.

Thereby, in this dissertation, the FastDTW algorithm proposed by Salvador S. *et al.* [91] was adopted. It is inspired in the constraints and data abstraction categories, resulting in a $O(s)$ time and space complexities. FastDTW uses a multilevel approach that recursively projects a solution from a coarse resolution and refines it. It shrinks a time series into a smaller time series, representing the same curve as accurately as possible but with fewer data points - Coarsening. It is run several times to originate different resolutions, then finds a minimum distance warp path at a lower resolution, using it as an initial guess for a higher resolution's minimum distance warp path - Projection. Lastly, it refines the warp path projected through local adjustments. In this approach, the cost matrix is only filled in the neighborhood of the path projected from the previous resolution, Figure 3.7. However, there is a risk that the optimal solution could not be found if the optimal path is not contained within the projected path. To overcome this, a radius parameter exists to control the additional number of cells on each side of the projected path that will also be evaluated in the refinement step.

3.4.1.2 Template Creation

The process of creating a sequence template consists on averaging a set of sequences representing the same type of activity. When working with time series data, a shape based averaging algorithm

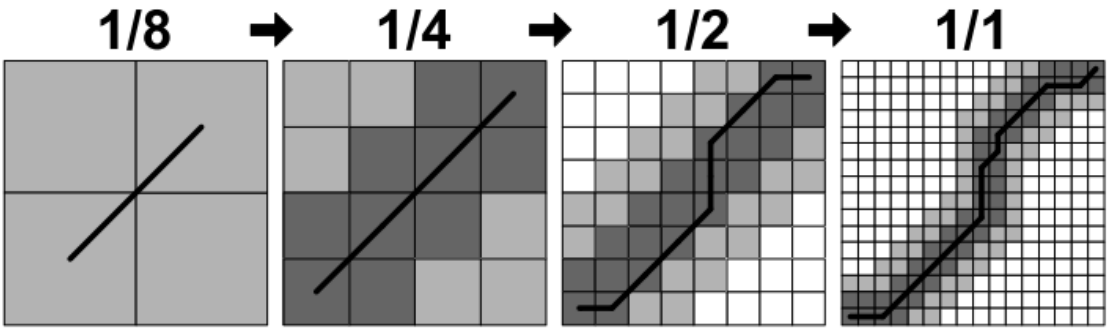

Figure 3.7: The four different resolutions evaluated during a complete run of the FastDTW algorithm, reproduced from [91].

is needed instead of a typical averaging approach, since in these sequences there is correlation among adjacent dimensions of time series [93].

In this work, a global averaging strategy called DTW Barycenter Averaging (DBA), introduced by Petitjean *et al.* [87], was used. DBA is a heuristic strategy, where an initial average sequence is iteratively refined in order to minimize the sum of squared DTW distances from that sequence to the set of existing sequences, Figure 3.9. In contrast with other averaging approaches presented in the literature, the DBA method is independent of the order the sequences contribute to the update. Being $S = \{S_1, ..., S_N\}$ the set of sequences to be averaged, $C = (C_1, ..., C_T)$ the average sequence at iteration i and $C' = (C'_1, ..., C'_t)$ the update at iteration $i+1$, for each refinement, the algorithm works as follows [87]:

1. DTW calculation - the DTW is computed between the temporary average sequence that will be refined and each individual sequence, to associate each coordinate of the average sequence to one or more coordinates of the sequences S, through an association function f.

2. Barycenter calculation - each coordinate t of the average sequence is computed as the barycenter of coordinates associated to it in the first step:

$$C'_t = barycenter(f(C_t)) \tag{3.3}$$

where,

$$barycenter\{X_1, ..., X_\alpha\} = \frac{X_1 + ... + X_\alpha}{\alpha} \tag{3.4}$$

$$X \in E$$

(E is a vector space)

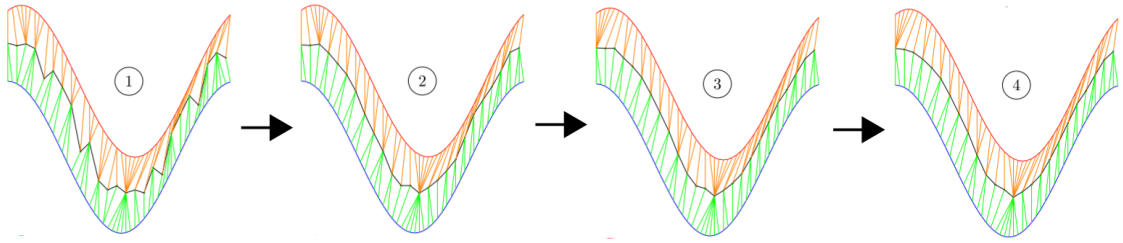

Figure 3.8: DBA iteratively adjusting the average of two sequences, reproduced from [87].

3.4.1.3 Implementation of similarity measures for FOG detection

The last two sections presented a brief description of the main concepts needed to understand the use of similarity measures. Now, the details concerning the implementation of these measures will be presented.

The similarity measures are applied to the four acceleration signals: acceleration modulus, horizontal forward, vertical and horizontal lateral accelerations. The extracted features are listed in Table 3.1. In order to compute them, three main tasks have to be completed: (i) Partitioning of the signal into gait cycles, (ii) Template creation and (iii) DTW computation.

Partitioning of the signal into gait cycles - needed to create the sequences that will be used in the DTW computation along with the template sequence. This task was already explained in Section 3.3.

Template creation - Since each person has a particular pattern of gait, a gait cycle template for each patient in the study needs to be created. The sequences used to create the template belong to the calibration signal, defined as the first ten consecutive cycles of normal gait, in the case of the data from the DAPHNet dataset. The gait cycles identification, through the process explained in Section 3.3, was performed using the acceleration modulus. However, the identification was transposed for the signal of each acceleration axis to have gait cycles of the four signals. Having defined the set of sequences to average, the template was created using the DTW Barycenter Averaging algorithm, described in Section 3.4.1.2, and available in the `tslearn` Python package [94], which provides machine learning tools for the analysis of time series, among them, the Python implementation of the DBA method.

The implementation of the method assumes that all the sequences have the same length, though a slightly variation on the length of the selected gait cycles may exist. Thus, previous to the averaging step, the cycles were linearly interpolated to have a number of samples equal to the mean size of the ten gait cycles. This operation was performed using the `interp1d` class available in `scipy.interpolate` module. Figure 3.9 schematizes the process of template creation using acceleration data of Patient 2.

DTW computation - At this point, the elements needed for the DTW computation are created. In the context of the FOG detection problem, the differences in the amplitude of the sequences being compared can be ignored. According to the analysis performed, this type of variations in the signal is not indicative of FOG and different normal gait cycles can have different amplitudes. To overcome the amplitude variations, the signal - the sequences and the template - was standardized according to Equation 3.5:

$$signal_{standardized} = \frac{signal - mean(signal)}{std(signal)} \tag{3.5}$$

Then, the DTW distance is computed using the `fastdtw` function, a Python implementation of the FastDTW algorithm. The distance provided by the algorithm is normalized dividing the result by the length of the originated warp path. Thus, the final similarity value varies between 0 and 1, Figure 3.9.

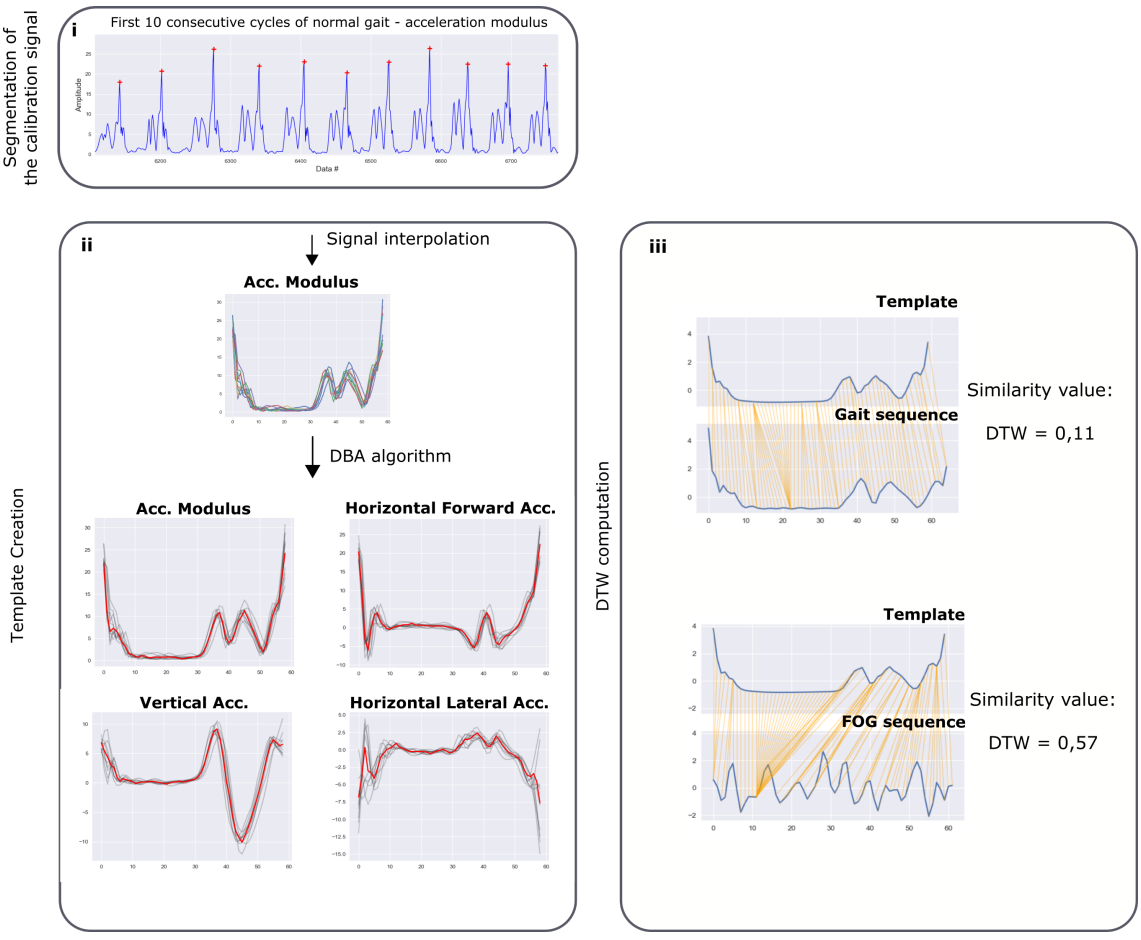

Figure 3.9: Main steps in the similarity measures implementation: (i) Segmentation of the calibration signal, using a peak detection approach, to create the sequences to be averaged, leading to a gait cycle template (ii). Then, the created template is compared with each window of the signal (iii).

3.4.2 Frequency - based Features

After the preliminary analysis performed on the signal, frequency-based features were also added to the set of features to be explored to solve the FOG detection problem.

From the signal analysis in the time and frequency domain, it was observed that for some patients, some freezing episodes are related to variations of the walking frequency. The literature also suggests that, for healthy subjects, during walking, most of the energy concentrates over [0,5 - 3] Hz - walking band. In contrast, for freezing events the energy is concentrated over [3 - 8] Hz - freezing band [62, 63, 71] . Figure 3.10 shows the amplitude spectrum of complete periods of gait and FOG, where the frequency variations could be observed, in contrast with the acceleration signal in time domain. It is important to notice that these frequency variations do not characterize all the freezing events of all patients, so an example where this feature is not clearly observed is also presented. In the normal gait periods one can observe that the higher harmonics are concentrated in the lower frequencies and in the first period of FOG there is a slight translation to higher frequencies.

3.4.3 Implementation of Frequency - based Features

In order to have representative data to computing frequency features, instead of compute them in a single gait cycle window, a set of three consecutive gait cycles was used, with an overlap of two gait cycles.

The frequency features, extracted for each one of the acceleration signal, are listed in Table 3.1:

To compute them the one-dimensional Discrete Fourier Transform (DFT) was implemented in Python, using the `fft` function available in the `numpy.fft` module.

Mathematically, being $x_1^W, ..., x_N^W$ the set of samples in a window W, the calculation of the DFT originates a set of complex values $X_1^W, ..., X_N^W$, each one representing the amplitude and phase of a harmonic, Equation 3.6:

$$X_h^W = \sum_{n=1}^{N} x_n e^{-i2\pi h \frac{n}{N}} \tag{3.6}$$

where $h = 1, ..., N$.

From the obtained result, the amplitude spectrum is computed for each window, and then the harmonics contained in the frequency band of [0.1-20] Hz are analyzed. The highest harmonic peaks are determined through a peak detector algorithm - `detect_peaks` function, available in [85].

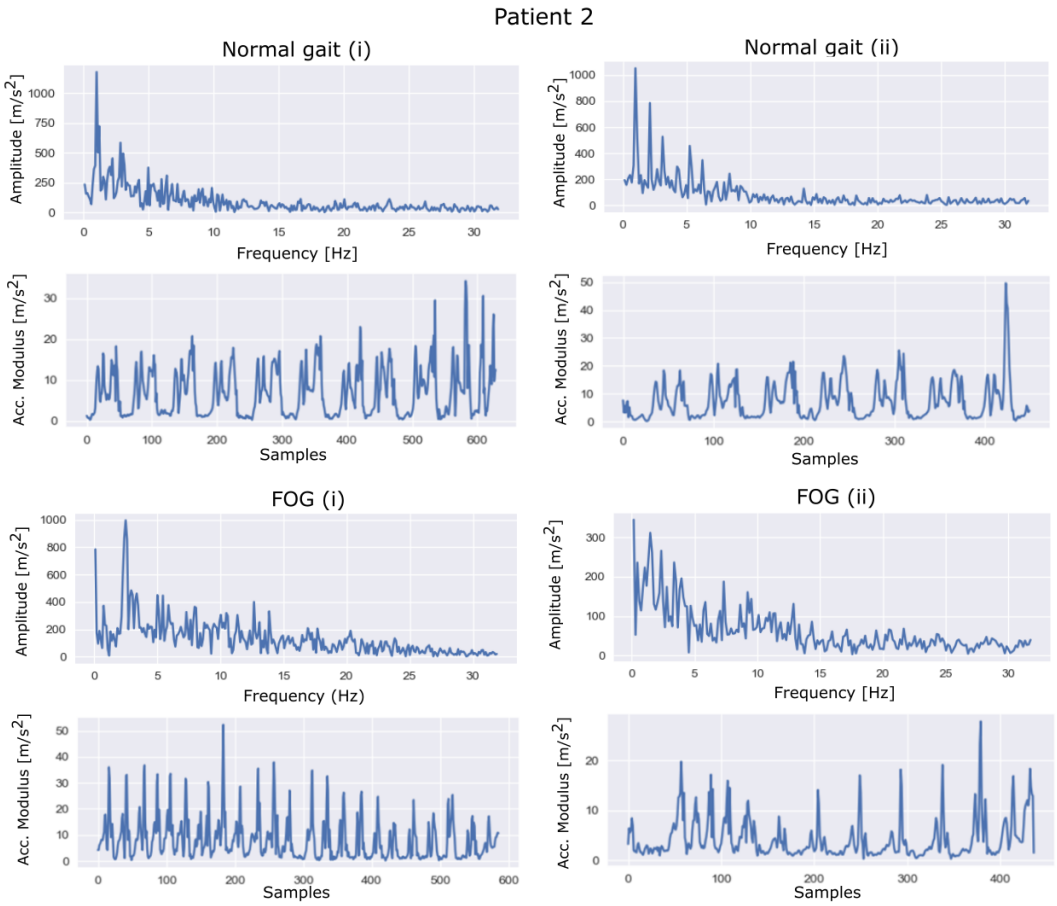

Figure 3.10: Normal gait and FOG events in the frequency and time domains. In Normal gait (i) and Normal gait (ii), the higher frequencies are concentrated in frequencies lower than in the FOG (i). In FOG (ii) this feature is not clearly observed.

Table 3.1: List of the features proposed in this dissertation. Their description and corresponding acronyms are presented.

Group of features	**Description**	**Acronym**			
		Acc. Modulus	**Horizontal Forward Acc.**	**Vertical Acc.**	**Horizontal Lateral Acc.**
Similarity Measures	DTW computation between a gait template and the acceleration signal sequences.	SM_M	SM_HFA	SM_VA	SM_HLA
Higher Harmonic Frequency	Frequency of the higher harmonic of the acceleration signal	HHF_M	HHF_HFA	HHF_VA	HHF_HLA
Higher Harmonic Amplitude	Amplitude of the higher harmonic of the acceleration signal	HHA_M	HHA_HFA	HHA_VA	HHA_HLA
Second Higher Harmonic Frequency	Frequency of the second harmonic of the acceleration signal	SHHF_M	SHHF_HFA	SHHF_VA	SHHF_HLA
Second Higher Harmonic Amplitude	Amplitude of the second harmonic of the acceleration signal	SHHA_M	SHHA_HFA	SHHA_VA	SHHA_HLA
Frequency Difference between harmonics	Difference between the frequencies of the two higher harmonics	DH_M	DH_HFA	DH_VA	DH_HLA

3.4.4 Traditional Features

Lastly, a group of features often used in activity recognition [86] and used by Mazilu *et al.* in [19] is also extracted. The goal is to test a traditional set of features in the created pipeline and compare its performance with the proposed features. In addition, in the Mazilu *et al.* work [19], the DAPHNet dataset was also used. Therefore, a direct comparison between the results obtained in this dissertation and the ones obtained by Mazilu *et al.* will be easier to perform.

The group of features, designated from now on as Mazilu's features, are summarized in Table 3.2. The first subset of features are extracted from each of the three accelerometer axis and the last six, using data from all the three axis. In total, 54 features are extracted from the acceleration data.

Table 3.2: Group of features proposed by Mazilu *et al.* and their description.

Features and Description	
Min	Minimum of the signal
Max	Maximum of the signal
Median	Median signal value
Mean	Average value of the signal
Root Mean Square (RMS)	Quadratic mean value of the signal
GeoMean	Geometric average of the signal
Variance	Square of the standard deviation
Standard Deviation (STD)	Mean deviation of the signal compared to the average
Kurtosis	The degree of peakedness of the sensor signal distribution
Skewness	The degree of asymmetry of the sensor signal distribution
TrimMean	Trimmed mean of the signal in the window
Entropy	Measure of the distribution of the frequency components
Asymmetry coefficient	The first moment of the data in the window divided by Standard deviation over the window
Range	The difference between the largest and smallest values of the signal
Zero Crossing Rate	Total number of times the signal changes from positive to negative or back, normalized by window length
Mean Crossing Rate	Total number of times the signal changes from below average to above average, normalized by window length
Signal Magnitude Vector (SMV)	Sum of the eucledian norm over the three axis over the entire window normalized by the window length. It indicates the degree of movement intensity, being very used in fall detection [95]
Normalized Signal Magnitude Area (SMA)	Acceleration magnitude summed over three axis normalized by the window length. It is used to distinguish between periods of activity and rest [95]
Average Acceleration Energy (AAE)	Mean value of the energy over three acceleration axis.
Eigenvalues of Dominant Directions (EVA)	Eigenvalues of the covariance matrix of the acceleration data along x, y and z axis.

3.4.5 Normalization and Mutual Information

After extracting the features representing each window of the signal, they are always standardized. They are rescaled to have the properties of a standard normal distribution with mean (μ) 0 and standard deviation (σ) 1. The standardization, also known as Z-score normalization, is required in

most of the machine learning algorithms and is computed according to Equation 3.7:

$$z = \frac{x - \mu}{\sigma} \tag{3.7}$$

Where x is the value of the extracted feature and z the rescaled value. At this point, having the array of extracted features and the corresponding labels, one can analyze the importance of knowing each feature to the labels identification, through the calculation of the Mutual Information (MI). The MI between two random variables is a non-negative value, that quantifies how much information the value of one variable reveals about the value of another. It is equal to zero if the two variables are independent and higher values means higher dependency [96]. Thus the MI of the features described above is analyzed and discussed in Chapter 5.

3.5 Classification

This section presents the main steps that are part of the classification procedures, along with a brief explanation of the main associated concepts.

3.5.1 Training and Test Procedures

With regard to the classification stage and consequent division into train and test sets, according to the literature review, one can identify two main approaches:

- Patient Independent - usually a Leave One Patient Out (LOPO) approach, where the model is trained with data of $N-1$ patients and then it is applied to the patient being tested.
- Patient Dependent - the model is trained only with data belonging to the patient being tested.

Using the first method, one has more data to train the classifier, creating a more generalized model, that would eliminate the need of acquiring labeled data from a test patient. However, as mentioned before, gait and FOG events are highly patient dependent and thus, there is a risk that the classifier will be trained with gait and FOG features that do not apply to all patients. On the other side, the Patient Dependent approach has the need of performing a first step of data acquisition and labeling with the data from the test patient. But, in this way, the created model would only capture the specificities of the patient gait and FOG, being more patient specific. Moreover, other studies have demonstrated that there are significant improvements when the algorithms are personalized, [18, 63, 64]. Therefore, a Patient Dependent approach is adopted in this dissertation.

The features described in Section 3.4 represent each signal window and are organized in an array of features that are the input of supervised learning classifiers. For each patient, the obtained array of features is divided in two different sets: a train and a test set. This division aims to use a train set to select the relevant features, through sequential feature selection algorithms and to tune and create the classifier, obtaining cross-validation metrics. Then, a test set is used to evaluate the model, in a new set of samples, never seen by the classifier.

The division is performed in a way that the train set comprises the first 75% of the labeled episodes, with the corresponding normal gait portions of the signal, and the test set comprises the remaining signal, namely, the last 25% of the labeled FOG episodes, with the corresponding no freezing episodes. The choice of including in the test set the last 25% of the FOG episodes, instead of simply define it as 25% of the totality of the signal is inspired on an approach presented in [18]. The freezing episodes distribution along the totality of the acquired signals varies from patient to patient. Thus, a blind division would create more unbalanced sets, concerning the amount of data of class FOG and class Gait. For example, a patient who did not manifest any FOG in the last part of the acquisition would have a test set with no event. In this way, a division based on the number of episodes is preferable.

Also, a fixed test set is always used. In other words, for each patient the last 25% of the episodes are always used to define the test set, instead of using a random set. This is done to get more direct comparisons between different models and techniques than the ones resulting of different random test sets.

3.5.2 Feature Selection

The array of features extracted for each patient is submitted to a process of sequential feature selection. It aims to select a subset of more relevant features for FOG detection. In this way, the computational efficiency is improved as well as the prediction performance, since the generalization error of the model is reduced by removing irrelevant features or noise [97].

A sequential feature selection algorithm is a greedy search algorithm that removes or adds one feature at a time based on the classifier performance until a feature subset of the desired size k is reached. Thus, an initial d-dimensional feature space is reduced to a k-dimensional feature subspace, where $k < d$ [98].

In this work, a Sequential Forward Selection (SFS) algorithm is implemented. The algorithm is initialized with an empty set of features. Then, the feature associated to the best classifier performance is added to the set. Instead of defining the k dimension of the new feature space, a $[min\_k, max\_k]$ range of features is defined. The SFS will then select the best feature combination that is discovered by iterating from $k = 1$ to max_k and return a subset with a size within max_k to min_k, depending on which combination returned best classifier performance during a 10-fold stratified Cross Validation [98]. The classifier performance evaluation is done using a criterion function, namely, the geometric mean (GM) between specificity and sensitivity, Equation 3.8. In this way, one intends to maximize both specificity (the proportion of negatives correctly identified) and sensitivity (the proportion of positives correctly identified).

$$GM = \sqrt{Specificity \times Sensitivity} \tag{3.8}$$

where,

$$Specificity = \frac{TrueNegatives}{TrueNegatives + FalsePositives} \tag{3.9}$$

$$Sensitivity = \frac{TruePositives}{TruePositives + FalseNegatives} \tag{3.10}$$

The implementation of the described algorithm is done using the `mlxtend`, a machine learning extensions Python library [98], and its `SequentialFeatureSelector` class [99].

The $[min\_k, max\_k]$ range of features is defined according to the set of features that is being tested. When one is only using the similarity measures as the input of the algorithm, the *min_k* is defined as 2 and the *max_k* is defined as 4. For sets with a higher number of features (the frequency features together with the similarity measures or with the Mazilu's features) the max_k is defined as 20, because it was verified that the performance of the classifiers stabilizes around that number of features.

3.5.3 Classifiers

In order to solve the FOG detection problem, different classifiers have already be used in related works. In this thesis, Random Forests (RF) and SVM were explored. A brief description of each one of these classifiers is now performed:

Random Forest - It is a non parametric algorithm and an example of an ensemble method built on decision trees, since it relies on aggregating the results of an ensemble of simpler estimators [100]. A simple decision tree (estimator) splits the samples along one or the other axis, according to some quantitative criterion. Then, each created region is labeled taking into account the class that is better represented in that area. If the created region has samples of more than one class, it is split again along one of the axis.

As the depth (number of splits performed) increases, the decision tree tends to overfit the data. A RF combines multiple decision trees to reduce this effect of overfitting. An ensemble of parallel random estimators fits various sub-samples of the dataset and averages the results to find a better classification. In this work, the classifier was implemented through the `RandomForestClassifier` class available in the `Scikit-Learn`. The number of estimators (number of trees in the forest) is the only parameter that needs to be defined.

Support Vector Machine - It is a supervised machine learning algorithm that separates two classes by a linear decision boundary - hyperplane - that maximizes the margin. In other words, it defines a hyperplane that has the largest distance to the nearest training data point of any class. These closer samples are called support vectors [101]. The problem can be formulated and solved in quadratic optimization terms, as follows:

$$min[\frac{1}{2}\|w\|^2 + C(\sum_{i=1}^{N}\zeta_i)] \tag{3.11}$$

given $y_i(w_0 + w_{z_i}^T) \geq 1 - \zeta_i, \forall i \in N$, defining the hyperplane and being y the vector of labels, z the input sample features and w the set of weights.

Sometimes the data set is not linearly separable. In these cases, one can either relax the hard margin by slack variables ζ_i, allowing misclassifications or use the *kernel trick* to map input vectors to a higher dimensional feature space where the linear separation may be possible. Thus, the typical hyperparameters of an SVM are the kernel function and C parameter, which trades-off misclassification of training examples against simplicity of the decision boundary. Thus, the decision surface is smooth with lower values of C, while higher values aim at correctly classify all the samples, giving the model freedom to select more samples as support vectors [102].

For the FOG detection problem, a Gaussian Radial Basis Function (RBF) is used as kernel function, due to its good performance and generalization capacity [18]. Consequently a new parameter needs to be adjusted, the γ parameter that defines how far the influence of a single training example reaches. When γ is low, the decision region is very broad and when it is high, decision-boundaries islands around data points are created, because of the their low influence [102]. The Python implementation of the Support Vector Machine was performed using the `SVC` class of the `sklearn.svm` available module.

By default, the models assume that the two considered classes are balanced - are equally represented in the dataset. But, in the FOG detection problem, it is not verified. The number of samples belonging to the Gait class is higher than the number of samples belonging to the FOG class. Additionally, the proportion Gait / FOG samples is also very variable from patient to patient. To deal with this unbalanced issue, both RF and SVM Python implementations have a `class_weight` parameter that allows to set different weights to different classes. Moreover, it has a `'balanced'` mode, which uses the values of the labels y to automatically adjust weights inversely proportional to class i frequencies in the input data, Equation 3.12.

$$class\_weight(i) = \frac{number\_of\_samples}{number\_of\_classes \times number\_of\_samples\_class(i)} \tag{3.12}$$

In the case of the RF, each sample will have a weight depending on the class it belongs, which will influence the way the splits and the classification of each region will be performed.

Concerning the SVM, the calculated weight is used to adjust the penalty parameter C, as $class\_weight(i) \times C$. Hence, since the number of class FOG samples is always much smaller, by Equation 3.12, the C parameter of this class will increase, meaning that one wants to put more emphasis on this class, to correctly classify all the samples.

Table 3.3: List of classifiers and parameters tested in the FOG detection problem

Classifier	Parameters
Random Forest	Number of estimators: {1, 3, 5, 10, 20, 30, 40, 50}
Support Vector Machine	C: {0.01, 0.1, 1, 10, 100, 1000, 10000, 15000, 20000, 50000} γ (RBF) : {10, 1, 0.1, 0.01, 0.001}

Each classifier is trained in the machine learning pipeline. For each patient, a set of hyperparameters is evaluated during the training phase, to find the ones that lead to the best performance of the

classifier for the concerned patient. In other words, a different model is created for each individual. The best parameters are found through an exhaustive search over a specified set of parameters values for an estimator. The exhaustive search is implemented using the `GridSearchCV` function available in `sklearn.modelselection` Python module. The parameters of the desired estimator are optimized by a stratified 10-fold cross validation grid-search over a predefined parameter grid - defines the set of parameters to be evaluated. Table 3.3 presents the list of classifiers and parameters tested. Additionally, to perform the parameters evaluation, a scoring metric needs to be defined. Similarly to what happened in the Feature Selection phase, the GM between specificity and sensitivity (Equation 3.8) is used.

3.5.4 Performance evaluation

The optimizations performed in the algorithm of FOG detection aim at the maximization of the GM. With this score function a join maximization of the sensitivity and specificity is achieved. GM varies between 0 and 1, with higher values representing better performances. Other metrics may also be utilized. The FOG detection problem is an imbalanced problem, because the classes are not represented equally: there are much more samples belonging to the Gait class than to the FOG class. Thus, the analysis of the traditional accuracy score is not reliable. In these cases, the F1-score, Equation 3.13, could give more insight into the accuracy of the model. So, this is another adopted metric that will be discussed in the evaluation and comparison of the methods of FOG detection.

$$F_1 - score = \frac{2TruePositives}{2TruePositives + FalseNegatives + FalsePositives} \tag{3.13}$$

3.6 Post Processing

The Post Processing step consists in analyzing the predicted values in the test set, resulting from the classification pipeline, and processing them to improve the classified output quality, by removing noise, for example.

Even though this step cannot be applied in real-time detection of FOG, it can be relevant in a post signal analysis context. Two strands of post processing can be applied to the classification output, that will be named as *gap filling* and *noise removal*. Let 1 be the output of the classifier representing FOG and 0 the output representing normal gait:

The *gap filling* consists in setting samples from small periods of normal gait (comprising one or two samples), detected between FOG episodes, as samples of a FOG event. This processing favors the detection of FOG periods and it is done because it is assumed that the patient does not go from a FOG state to a normal gait state of one or two gait cycles and then to a FOG state again. This small period is then considered as FOG, as well. In a simple way what happen is: $0000011011110000 \rightarrow 0000011111110000$.

The second type of processing, *noise removal*, is the action of ignoring punctual detection of FOG, considered as noise. This is, ignoring isolated samples that are detected as FOG and considering it as normal gait - $00100000011110 \rightarrow 00000000011110$. An isolated detection means that a FOG event, with a gait cycle duration, is detected. This will be a very small period of blocking (about 1 second), with a very low clinical relevance. Actually, some authors defend that the minimum length of an episode with clinical importance must be 3 seconds [103].

Chapter 4

A first step into pre-FOG detection

Inspired in a few works that explored the existence of a pre-FOG phase - a period when the characteristics of gait start to degrade before a true episode of freezing occurs - in this work a simple analysis of pre-FOG phase is also performed. These studies are based on the threshold model of FOG, mentioned in Section 2.1.2, which suggests that an accumulation of motor deficits until a threshold is reached leads to a FOG [37].

The analysis adopted in this dissertation is based on the study of Palmerini *et al.* [5] and intends to compare, for each individual, the behavior of the similarity measures and of the frequency features, described in Section 3.4, during normal gait and during a few moments before the beginning of a FOG. The same procedures for signal segmentation and feature extraction, explained in the previous chapter, are implemented here. The comparison between the two periods is performed using concepts of statistical analysis, namely, using a paired t-test in order to evaluate whether there is a statistical evidence that the difference between the features during pre-FOG and gait are significant. There are two main steps, necessary to accomplish the proposed analysis: (i) Identification of the pre-FOG and gait sequences before each FOG and (ii) Statistical analysis of the features of these two sets of data. The following sections detail these two steps.

4.1 Identification of the pre-FOG and gait windows

In the pre-FOG analysis, the DAPHNet dataset continued to be used. Since there is only information about the normal gait and the FOG episodes, a time period before each labeled FOG needs to be defined as pre-FOG. A period of 3 gait cycles (3 windows) before each FOG is then chosen. This period was chosen having in mind that it is a parameter that could be highly patient dependent, but also, variable with the context where each FOG occurs. However, 3 gait cycles were assumed to represent an appropriate period to explore the existence of gait deterioration before a FOG.

For each patient, sets of gait and pre-FOG windows, manifested before each identified FOG, will be part of the statistical analysis. During the creation of these sets, two particular situations can happen:

- the pre-FOG period overlaps a FOG event, that is, there are not 3 gait cycles before a FOG;
- the number of gait windows is smaller than the number of pre-FOG windows (smaller than 3).

In these cases, the sets of data are discarded and are not part of the statistical analysis. Figure 4.1 reflects these situations.

Additionally, since the pre-FOG phase is only associated with FOG episodes that take place during movement, gait windows that may represent periods of inactivity - when the patient is stopped - are also ignored. These periods are not labeled in the dataset, but are easily recognized in the acceleration signal. A threshold-based approach, applied in some extracted features, can be used to discriminate between presence and absence of motion. Hence, the variance - which measures the variability of a distribution and is mathematically defined as the average of the squared differences from the mean - can be chosen for this purpose [104]. From an analysis performed on this feature behavior during gait periods, in data from all the patients, a threshold of $5m^2/s^4$ was experimentally defined to discriminate the two moments. Hence, windows with a variance smaller than the threshold are considered windows with insufficient motion and are then eliminated.

Having the sets of gait and pre-FOG windows, the average of the extracted feature in these two types of windows are calculated, Figure 4.1. Hence, each pair (gait - pre-FOG) is associated to an individual FOG of a specific subject, where the first value is the average of the feature values of the gait windows and the second is the average of the feature values of the pre-FOG window.

4.2 Statistical Analysis

After computing every (gait - pre-FOG) pair, one is now able to do the statistical analysis. This analysis is performed separately for each one of the eight freezing patients present in the DAPHNet dataset and for each one of the features, considering each FOG as a single condition of FOG. In the work of Palmerini *et al.* [5], the pairs related to FOGs of the same condition (the conditions are, for example, turning, walking straight, walking in circles) were aggregated in a way that for each patient and for each condition a pair (gait - pre-FOG) was created. However, in the present work there is no information about the context of the events to perform this differentiation. Moreover, their analysis is done for each feature, with data from all patients, instead of carrying out a patient dependent analysis, as suggested in the present work. Since the FOG and consequently the pre-FOG events are highly patient dependent, a patient dependent analysis is here considered a better approach.

The comparison between features in the two periods is performed through a paired t-test, implemented in Python through the `ttest_rel` statistical function available in the `scipy.stats` module. The level of significance p was set as $p = 0,05$. A paired t-test (or dependent sample t-test) evaluates the significant difference between two related variables, for example, before and after observations on the same subject. The variables in study in this work, the features during gait and pre-FOG, are considered related variables, and thus, a paired t-test is suitable to be used. The

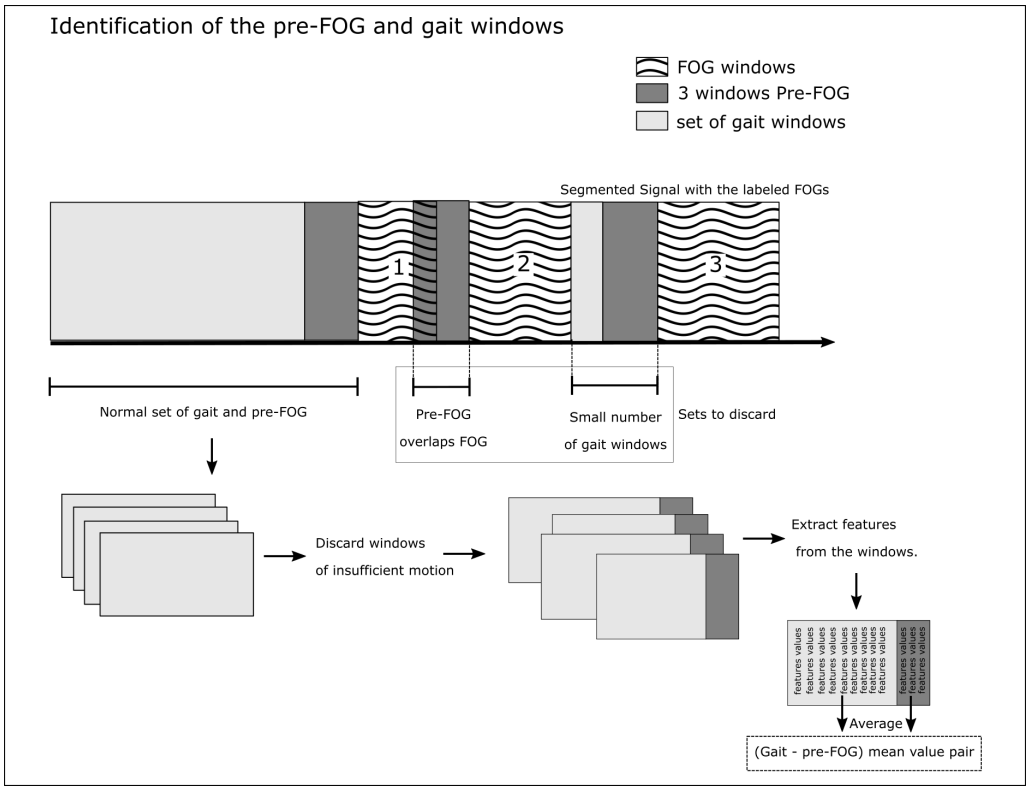

Figure 4.1: Schema of the steps involved in the pre-FOG and gait windows identification and the determination of the (gait - pre-FOG) pair.

null - hypothesis assumes that the true mean difference between the paired samples is zero and it is rejected if the result of the statistical test is lower than the level of significance [105].

The paired *t*-test assumes that the dependent variables are normally distributed. The testing for normality should be conducted on the differences between the two conditions and can be performed using the *Shapiro-Wilk* test, which tests the null-hypothesis that the data is drawn from a normal distribution. Thus, this test is implemented before the paired *t*-test and, in the case the result indicates that the variables are not normally distributed, the *Wilcoxon signed-rank* test is applied instead. It is also used to compare two related samples and it is usually applied as an alternative of the paired *t*-test when the population cannot be assumed to be normally distributed.

Additionally, the *p* values resulting from the statistical tests mentioned above need to be corrected, since a multiple testing procedure is performed. A multiple test arises when a statistical analysis involves multiple simultaneous statistical tests. In this case, for each patient the number of tests performed is equal to the number of features tested. There are a lot of methods that could be applied to perform the correction of the results. Here, the same procedure adopted in [5] was used, the Benjamini and Yekuteli procedure [106]. Thus, the significance of the results is only assessed after recalculating the probabilities obtained from the statistical tests.

Chapter 5

Results and Discussion

The methods presented in the last two sections were implemented, leading to a set of outcomes. The results obtained in the main phases of the algorithm are presented in this chapter. They concern the analysis and evaluation of the extracted features, the algorithm on FOG detection performance and comparison with literature results and, lastly, the results of the analysis of the periods before a freezing event - pre-FOG analysis. Throughout the results presentation, they are also discussed. At the end, a few remarks about the computation cost and the real-time implementation are presented. The results were obtained using only inertial data from the Daphnet dataset, contrarily to the initial goal of the dissertation where the acquisition of data from a new set of patients was expected.

5.1 Feature Extraction

The feature extraction phase is one of the most important phases in the entire pipeline. The extraction of a representative and significant set of features, with relevance for the detection problem in hands, is the starting point to achieve good results. Notice that one of the concerns of this work was to find meaningful features, with particular interest in the FOG detection, instead of using a set of generic features of signal processing and activity recognition, since the freezing events represent a special type of activity. Hence, the analysis of the results obtained from the feature extraction is important to evaluate their relevance. The following sections present these results.

5.1.1 Similarity Measures

This section provides the results of the similarity measures extraction through the DTW distance computation. Figure 5.1 presents the DTW result in a window corresponding of normal gait and in a window of FOG, belonging to Patient 1, using the four signals: acceleration modulus, horizontal forward, vertical and horizontal lateral accelerations.

From the obtained values in the illustrative windows, one can infer that these features have a potential to distinguish between FOG and normal gait, since the FOG window has DTW values higher than the normal gait window. It is an expected result since the FOG episodes correspond

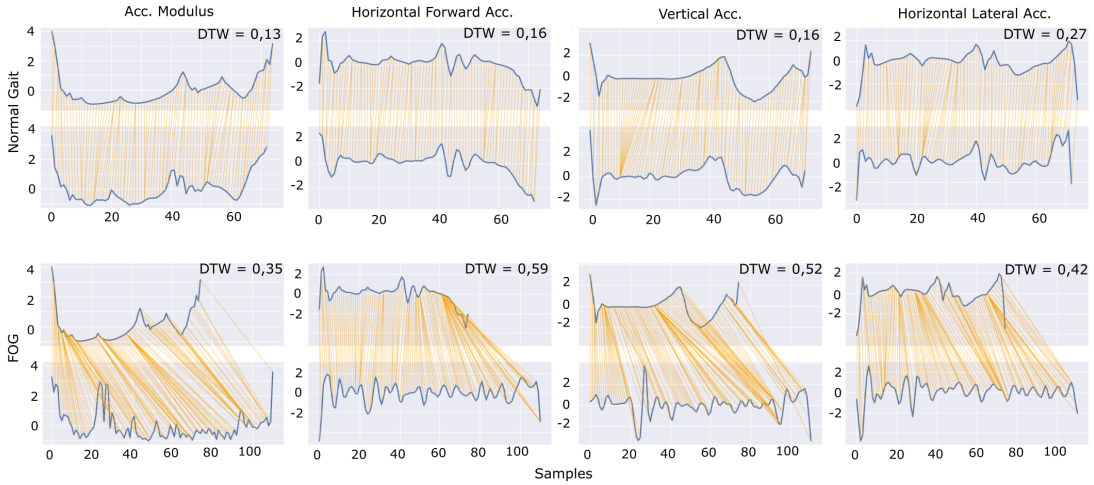

Figure 5.1: Example of the similarity measures extraction by computing the DTW in windows of normal gait and of FOG. The computation is illustrated for the four used signals. The vertical axis are the amplitude of the standardized signals and the horizontal axis are the samples. The value of DTW varies between 0 and 1.

to degradation of gait. However, this is just a pair of normal gait - FOG window for each signal. Figure 5.2 presents the DTW computation result in a bigger set of windows, corresponding to almost 3 minutes of the recorded signal, for Patient 1 and Patient 2. The acceleration modulus signal, that appears in the background of the graphics, was used to compute the DTW measure. The freezing events are identified by the red line and the value of the distance by the green line. As one can see, mainly in Patient 2, the FOG events are associated with an increase of the distance value, highlighting the discriminative power of this feature. In Patient 1, this behavior is also possible to be observed, mainly in the longest episode. The remaining episodes are smaller and thus, this effect is not visible, maybe because, in such a rapid event, the gait degradation is not so intense inducing smaller changes in the acceleration signal.

Another important behavior manifested in both graphs is the DTW values in periods of absence of movement (for example, the period between the 18000 - 20000 samples in Patient 1, when the acceleration signal is near 0 m/s^2). Here, the pattern of the signal is even more different from gait than the pattern of FOG, leading to DTW values much higher than the remaining.

In summary, based on this analysis, the extracted similarity measures seems to have a good discriminative power to distinguish between FOG and gait. The use of these features is advantageous because they are adapted to each patient, as a template needs to be created for each individual. Additionally, it is a simple measure, with an intuitive meaning, making easier to understand its behavior.

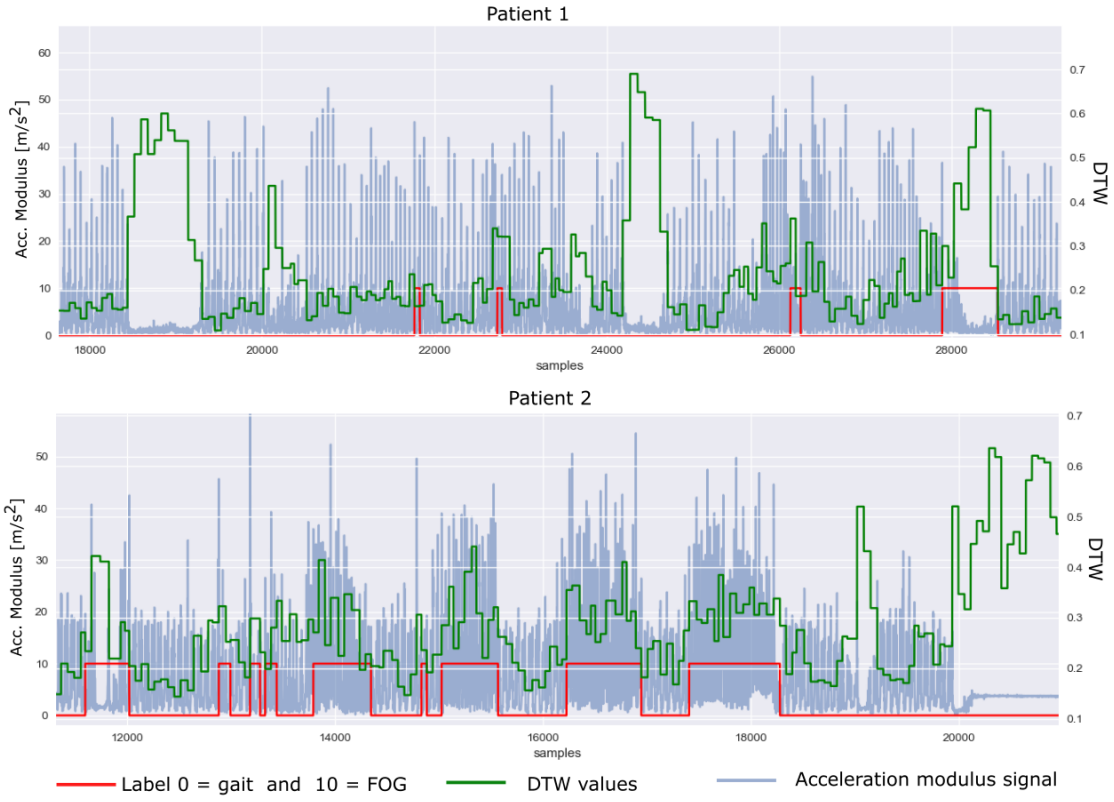

Figure 5.2: Behavior of the DTW similarity measure computed using the acceleration modulus signal throughout approximately 3 minutes of recorded signal, from Patient 1 and Patient 2. The values were computed for each window and are reported in parallel with the acceleration signal and the labels.

5.1.2 Frequency-based Features

Similarly to the analysis performed above, the frequency-based features values extracted in a set of gait and FOG windows are also reported in Figure 5.3. Once more, the presented features were extracted from the acceleration modulus signal, represented on the background. Notice that the majority of the examples are presented with this signal since it can encompass the main characteristics of each acceleration axis, avoiding the need of presenting an example with all the used signals. Nevertheless the remaining signals still have their relevance. In this case, each graphic represents approximately 1,5 minutes of recording.

Concerning the frequencies of the higher and the second higher harmonic, a slight increase in their values is observed when a FOG window takes place. These differences are more evident in some FOG than in others - each FOG is particular. Moreover, the amplitude of the same harmonics have a behavior that follows the amplitude variation of the acceleration signal itself, which, sometimes, can be significant to distinguish between FOG and gait. Lastly, the difference between the frequencies of the first and the second harmonic seems to be the frequency feature with less potential to identify the two classes.

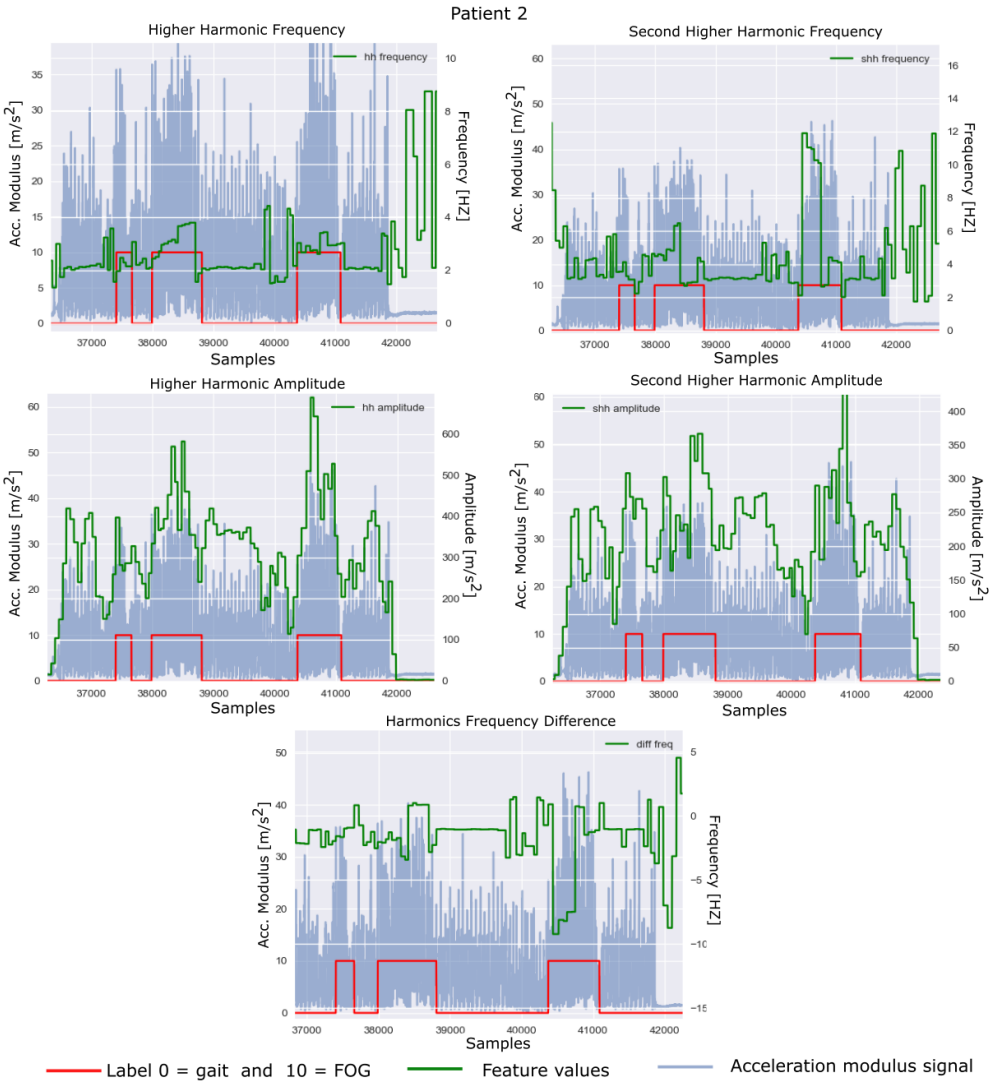

Figure 5.3: Behavior of the frequency-based features throughout approximately 1,5 minutes of recorded signal, from Patient 2. The values were computed for each window and are reported in parallel with the acceleration signal and the labels.

Both in the similarity measures presented before and in these features, even through the signal corresponding to gait periods, there are many moments where considerable changes in the features values occur, but they do not mean that the patient is facing a FOG. This is one of the reasons for the high number of false positives detected in the classification step. What happened during these moments is impossible to know, since one only has access to the labeled data. The cause of these variations would be interesting to explore had we our own data acquisitions.

5.1.3 Mutual Information of the Features

The extracted features can also be evaluated by computing the MI. As mentioned in Section 3.4.5, the MI is a measure of the mutual dependence between two variables. Thus, the MI was computed for each feature against the labels. Figure 5.4 presents the MI values for all the features explored in the dissertation, for each patient of the DAPHNet dataset. The first four features correspond to the similarity measures, the next twenty features concern the frequency-based features and the remaining ones to the Mazilu's features [19]. The green line represents the mean of the MI for the two main sets of features: the proposed in this dissertation and those proposed by Mazilu *et al.* [19].

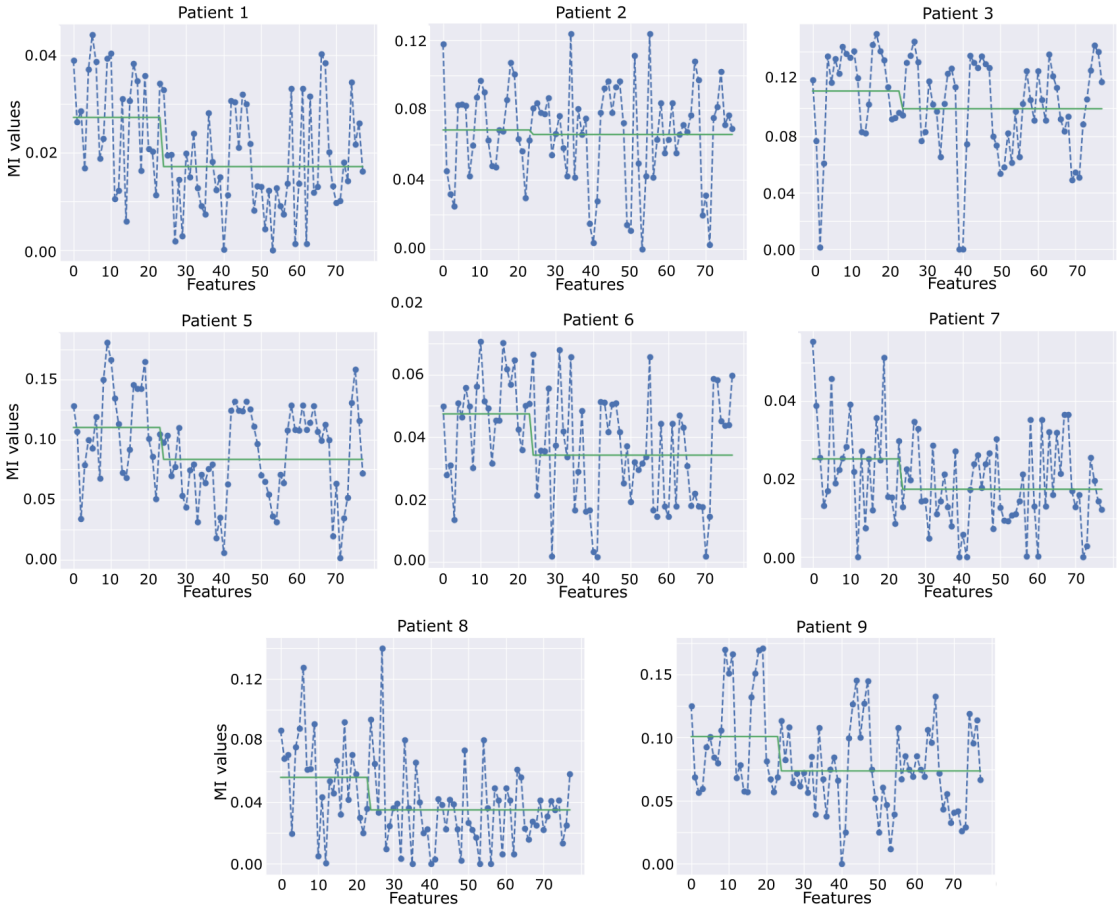

Figure 5.4: Features MI values for the eight patients of the Daphnet dataset. The line represents the MI average of the two main sets of considered features, the proposed in this dissertation (similarity measures and frequency-based features) and ones used by Mazilu's

First, as expected, one can see that same features have different values of MI for different patients and the best features of one patient are not necessarily the best features of another, highlighting once again the patient dependent nature of the freezing episodes. Each patient has a particular way of manifesting FOG, which is then characterized by different sets of features. For

example, for the Patient 7, the feature 0, corresponding to the SM_M, is the feature with the higher value of MI. For the Patient 6, this feature has a relative high value of MI, however there are many others with higher values.

Some common aspects can also be visualized. Among the four similarity measures, the SM_M is always the one with a higher value of MI, and for five of the eight patients tested, the SM_HLA (feature 3) is the one with poor values. The last is an expected result, since this acceleration axis is the one with less information about gait. Feature number 40, corresponding to the geometric mean of the vertical acceleration signal, appears in all the eight patients as a feature with low values of MI, evidencing that maybe this feature could be removed with no significant losses. Additionally, the set of features from 42-47, corresponding to the features of variance and standard deviation of the three acceleration axis, has always relatively high values of MI, meaning that they may have a valuable contribution to distinguish between FOG and gait.

The relative variation of the frequency-based feature values is also quite different across the patients, with no significant aspects to point out.

The similarity measures and the frequency features seem to have, in general, relative good MI values. In fact, for all the patients, the left side of the graph reports higher values of MI. This tendency is better visualized with the MI average for the two considered sets. As one can see, the MI average of the proposed features is always higher than the MI average of the Mazilu's features [19]. This result shows that the proposed features have a good potential to distinguish the two classes (FOG and gait), with better results than those proposed in the literature.

5.2 Feature Selection

As mentioned in Section 3.5.2, the set of extracted features is the input of supervised learning algorithms and the step of feature selection is done based on the performance of the chosen estimator. Thus, the features selected in this step can give insights on the importance of each input feature. To evaluate this, two scenarios are analyzed:

- the set of selected features, for each patient, when the input comprises the similarity measures and the frequency-based features.
- the set of selected features, for each patient, when the input comprises the similarity measures, the frequency-based features and the Mazilu's features.

With the results obtained from these scenarios, the graphic of the Figure 5.5 was created, reflecting the number of times (among the eight patients) each proposed feature was selected to proceed in the machine learning pipeline. The estimator used to obtain these results was the SVM.

When the input comprises only the proposed features, the ones that are more frequently selected are SM_M and HHA_VA, chosen 6 times out of 8, HHA_M and SHHA_M, chosen 5 times out of 8, and HHF_M, HHF_VA and SHHA_VA, selected 4 times.

From this, one can see that the features from the acceleration modulus and from the vertical acceleration are more frequently selected, than those extracted from the remaining signals. It

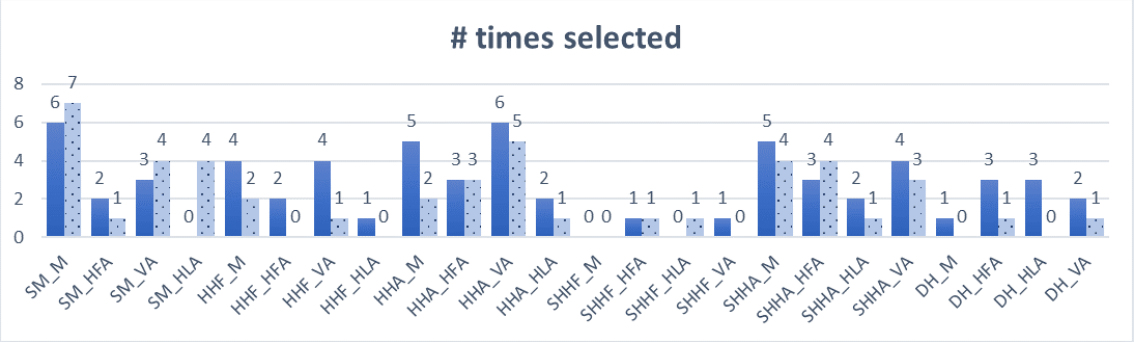

Figure 5.5: Feature Selection results, reporting the number of times each feature was selected across the eight patients. The solid bars report the results of the first scenario, where the input features are the similarity measures and the frequency-based features. The dotted bars represent the results of the second scenario, when the Mazilu's features are added to the input set.

shows that the vertical acceleration is the one that reflects more the characteristics of freezing, having then more information about these events. Extracting both the frequency and the amplitude of the higher harmonics was found to be advantageous, since both types of features are usually selected.

The fact that SM_M was selected more times than the other three, matches the results obtained from the analysis of its MI, in Section 5.1.3, since, as mentioned, this was always the similarity feature with higher MI values.

However, three features were never selected by the algorithm: SM_HLA, SHHF_M and SHHF_HLA. Once again, the low informative content of the horizontal lateral axis is confirmed.

Concerning the second scenario, when the Mazilu's features are added to the set, the results show that the proposed features still have a good representation among the selected ones. Even with the addition of more 54 features, the proposed ones are still selected to contribute to the classification problem. Moreover, the similarity measures, mainly the SM_M, are even better represented. In fact, the features based on the similarity measures are selected in all of the eight tested patients, demonstrating the reliability in the introduction of this new type of measures. In both sets, there are a few that are never selected, meaning that they never produce an increase of the classifier performance and thus, their computation could be unnecessary. It should be noted however that the algorithm was only tested in eight patients and due to variability in the FOG manifestation, making strong conclusions about the no need of a feature is very hard, since it could be irrelevant to a set of patients, but relevant for another one.

5.3 Classification results

This subsection performs the analysis of the results concerning the classification step. A comparison between the performance of the RF and the SVM in the FOG detection is performed, but

a greater focus is given to the SVM performance using different inputs, since it was the classifier with the best results, as presented in Section 5.3.1. The presented values are obtained before the application of signal post-processing for noise removal, in order to have a reliable analysis of the false positives detection. Only the *gap filling* processing was applied. Finally, Section 5.3.3 presents a detailed analysis of the prediction results obtained, comparing the predicted labels to the real ones and even to the raw acceleration signal to better understand where and why the algorithm performed not so well.

5.3.1 Different classifiers performance

Table 5.1 reports the stratified 10-fold cross validation results with the training set of the eight considered patients, using a RF and a SVM. The similarity measures and the frequency-based features were the input of both classifiers. The GM, which was the metric to maximize, is presented, as well as the specificity, the sensitivity and the F1-score.

Observing the values, the first aspect to point out is that, for all the patients, the SVM leads always to better values of GM, and thus, one can say that the SVM has better performances in the FOG events detection, than the RF. In fact, the average results show an increase in the GM from $66,0 \pm 12,3\%$ with the RF, to $83,9 \pm 6,0\%$ with the SVM, representing an increase of almost 18% in the FOG detection performance. This increase on the GM values is a result of the increase of positive samples detection (FOG samples), that is, of the sensitivity (from $50,8 \pm 5,8\%$ to $85,1 \pm 15,2\%$), with a small decrease in the specificity, meaning that less true negatives are detected - more gait samples are detected as FOG samples (false positives). The decrease in the specificity is a price to pay to have higher values of GM, since, in this context, one wants to favor the detection of FOG, even if it means that some events are detected when they are not actually happening.

The F1-score has also a small increase in its value with the SVM classifier, though it remains relatively low. The F1-score, defined by Equation 3.13, gives more weight to the true positive detections, but helps in the analysis of both false positives and false negatives. Consequently, with the high values of sensitivity achieved with the SVM classifier, these low values of F1-score mean that a high number of false positives are being detected.

With the results presented in the table, an individual analysis of the classifiers performance for each patient can also be performed. It is possible to see that the classifiers performance varies across the patients, with the RF having more variations than the SVM (standard deviation of 12,3 against 6,0). Focusing on the SVM results, the best performance was achieved with Patient 2, with 92,7% of GM and 81,2% of F1-score. On the other hand, the lower performance was achieved with Patient 8, with a GM of 71,6% and F1-score of 44,7%.

In short, the SVM is the classifier that leads to more consistent and higher performances in the FOG detection and thus it is the classifier one chose to use.

Table 5.1: Classification results obtained from the 10-fold Cross Validation performed with the training set of each freezer patient of the DAPHNet dataset, using RF and SVM. The values are in percentage.

	RF				SVM			
Patient	**Specificity**	**Sensitivity**	**F1-score**	**GM**	**Specificity**	**Sensitivity**	**F1-score**	**GM**
1	98,9	39,0	46,8	59,8	81,1	89,5	36,9	84,8
2	94,9	72,2	72,9	81,8	91,3	94,3	81,2	92,7
3	93,4	50,6	55,8	68,0	80,0	89,5	67,5	84,1
5	88,1	56,7	59,2	70,0	78,3	85,6	70,7	81,6
6	92,1	42,0	42,9	60,9	80,4	89,3	58,1	84,4
7	96,9	25,3	27,1	43,2	74,6	95,0	35,6	83,4
8	84,7	49,9	47,0	63,7	82,8	48,3	44,7	71,6
9	93,0	70,5	71,6	80,3	87,8	89,5	76,7	88,3
Average	**92,7**	**50,8**	**52,9**	**66,0**	**82,0**	**85,1**	**58,9**	**83,9**
STD	**4,6**	**15,8**	**15,3**	**12,3**	**5,3**	**15,2**	**18,0**	**6,0**

5.3.2 Different feature sets performance

In this section, the results obtained with the SVM classifier in the train and test sets (see Section 3.5.1) are reported, for different sets of features as input:

- Similarity measures
- Similarity measures + Frequency-based features
- Mazilu's features
- Mazilu's features + Similarity measures + Frequency-based features

Figures 5.6 and 5.7 show these results for the train and test sets, respectively. The presented values represent the average of the results obtained for the considered set of patients.

Concerning the proposed features, one can see that the four features of similarity alone have a relatively good performance, both in the train and test sets, evidencing the great potential of this kind of features to detect freezing episodes. Comparatively to the set of 54 features used by Mazilu *et al.* in [19], the performance of the similarity measures regarding the GM is only 7% lower in the training set and 3% lower in the test set.

The addition of the frequency-based features to the set with the similarity measures increased the GM results from $78,5 \pm 6,9\%$ to $83,9 \pm 6,0\%$ in the training set and from $77,8 \pm 6,3\%$ to $80,3 \pm 4,8$ in the test set. This means that the frequency-based features proposed have additional useful information that helps the model to decide better between FOG and gait.

The results of using the Mazilu's features in the proposed pipeline continue to be slightly higher. It is not necessarily a bad result. In fact, with a smaller and completely different set of features, some of them never used in the FOG detection context, the results are lower, but very close to the ones obtained with features used in literature. Moreover, joining the three sets of features leads also to an improvement up to 2% in the detection performance.

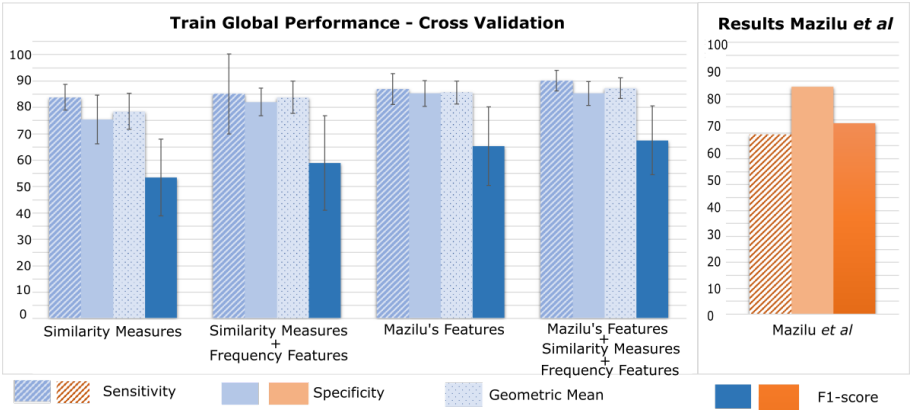

Figure 5.6: Global performance of the algorithm during cross-validation according to the set of features used as the classifier input. The metrics values are in percentage. In the right side, the results obtain by Mazilu *et al.* [19] with their methodology is also presented for comparison purposes.

The performances in the test set are slightly lower, resulting in differences up to 6% of the GM. The differences are not worrisome, since the results of cross-validation represent average results obtained in different folds and thus, it is expected that the model performs better in some folds than in others, which can be transposed for the test set. The test results on the Patient 6 are not included in the presented results. This patient was considered an outlier, since the model has very low performances in distinguishing between FOG and gait. The reason behind this could be a higher variety in the manifestation of freezing episodes. In the following section, a more detail analysis of this patient prediction results is presented.

In general the values obtained for the GM are relatively high, starting in 78,5 % with the similarity measures alone and achieving 87,3% with the 3 sets of features, in the train set. These values of GM represent a good balance between the values of specificity and sensitivity that are also good. On the contrary, the values of F1-score are relatively low - below 70% in the train set and below 50% in the test set. These values highlight an already mentioned issue: the detection of a high number of false positives. Many samples of normal gait are misclassified as FOG samples. The F1-score reflects the quantity of false positives in a more transparent way than the specificity and sensitivity, or even the GM. Notice that the metric that depends on the false positives is the specificity, Equation 3.9. However, having an unbalanced dataset, with more samples of class gait (negative class), sometimes the presence of many false positives can be overshadowed by the higher number of true negatives. In the literature, most of the works do not point out this issue, presenting the classification performance based only on metrics of sensitivity and specificity.

At this point, the results obtained with the proposed pipeline and sets of features, can be compared with the results presented in the literature. The first comparison that makes sense is the comparison of these results with the ones obtained by Mazilu *et al.* [19] with their features and pipeline, reported in Figure 5.6. Observing the values, one can see that their results of sensitivity

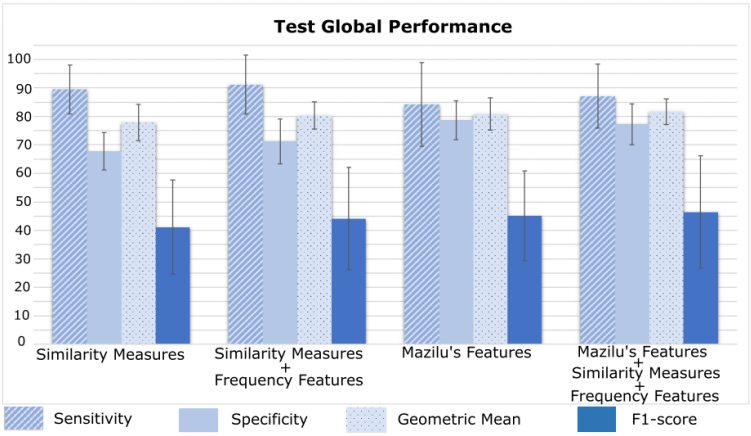

Figure 5.7: Global performance of the algorithm in the test set according to the set of features used as the classifier input. The metrics values are in percentage.

are relatively low, meaning that from all the FOG samples, only 69,4% were identified as such. For all the sets of features tested, including the Mazilu's features, the sensitivity obtained with the proposed methodology is always higher than 80%. On the contrary, their specificity and F1-score values are slightly higher than those presented here. This can be due to the fact that, in their pipeline, they removed portions of samples from pre-FOG periods. That is, the samples concerning a few seconds before the starting of a FOG were removed and were not part of the training and classification processes. These samples belong to a critical portion of the signal, since delays and uncertainties in the identification of the exact starting of a freezing event may lead to an erroneous labeling of these samples. Moreover, in some patients, these periods may correspond to a pre-FOG phase, where the gait pattern is already degrading. Consequently, more false detections are expected to occur in these periods of the signal. If these samples are removed from the problem, the performance of the classification may then improve, which can justify the results obtained by Mazilu *et al.* [19].

The results can also be compared to other studies related to the detection of FOG and already presented in Section 2.3.1 of the Background and Literature Review. Table 5.2 reports these results and also the dataset used in each work. It can be seen that the results obtained with the approach presented in this dissertation are better than the ones presented in a few studies, namely in the study of Bachlin *et al.* [63] and Moore *et al.* [57]. Although recent studies reported better performances, some of them with values of sensitivity and specificity higher than 90%. A few aspects can explain these results. In the other works of Mazilu *et al.*, [64] and [8], acceleration signals from different locations are used simultaneously to performe FOG detection. In the former, the signal from the three locations available in the DAPHNet were used, namely from the lower back, shank and thigh. In the latter, they have used signals from the two ankles. Hence, more information about the freezing events is given to the model, which could lead to better detection results. In the presented approach, only the acceleration signal from an ankle was used and thus, it could be one

Table 5.2: Reported results in literature concerning FOG detection. The values are in percentage.

Work	Dataset	Sensitivity	Specificity
Bachlin *et al.*, 2009 [63]	DAPHNet	73,1	81,6
Mazilu *et al.*, 2012 [64]	DAPHNet	98,4	99,7
Moore *et al.*, 2013 [57]	25 PD patients	84,3	78,4
Mazilu *et al.*, 2014 [8]	5 PD patients	97	-
Ahrichs *et al.*, 2015 [65]	REMPARK project	100	91,1
Rodriguez-Martín *et al.*, 2017 [18]	REMPARK project	88,1	80,1
Lorenzi *et al.*, 2017 [66]	16 PD patients	94,5	96,7
Proposed approach	DAPHNet	85,1	82,0

of the reasons for the differences in the performances. Ahrichs *et al.* [65] reported very high values of sensitivity and specificity. Their approach is based on an aggregation of the classifier's outputs to determine a confidence value, which is used for the final classification. Without this post processing computation, their sensitivity decreases in approximately 5%, being closer to the achieved by the presented approach. The work of Rodriguez-Martín *et al.* [18] reports results very similar to the ones presented. Their sensitivity is just 3% higher and, on the contrary, the specificity is 2% lower. Their results also concern a patient dependent approach, but they used data from the test patient together with data from the remaining, which can increase the variability of FOG events that the model learns. Finally, the work of Lorenzi *et al.* [66] clearly outperforms the results obtained. In their work, in addition to using the acceleration signals from sensors placed on both shins, they also used the signals from the gyroscope. Moreover, the latter works have used a dataset different from the one used in this dissertation, which could lead to variations in the results.

In short, the results of the proposed approach are well positioned among the results presented in the literature, evidencing the reliability of the method in the FOG detection.

5.3.3 Analysis of the test results

In this section, the prediction results obtained with the test set of each patient and using the similarity measures and the frequency-based features as input of the classifier are presented. The main goal is to analyze the classification performance in a more intuitive way than the comparison of the values of the metrics. This way, one tries to simulate what would happen if the patient was walking and a system was monitoring the signals.

Table 5.3 reports these results and Figure 5.8 illustrates them. The real labels of the samples are also shown in the figure. Making a general analysis, it can be seen that the quantity and distribution of the episodes and their duration are different across the patients, as expected. There are patients experiencing just a few long episodes during the test set, such as Patient 8, and others, as Patient 5, experiencing many small episodes. Once again, this is due to the patient dependent nature of the freezing events and one of the reasons to adopt a patient dependent approach in this

Table 5.3: Prediction results obtained with the test set of each patient, before and after the *noise removal* post processing step. The latter are presented inside brackets. The values are in percentage.

	Test Metrics (Post - Processing Test Metrics)			
Patient	**Specificity**	**Sensitivity**	**F1-score**	**GM**
1	72,5 (75,9)	90,9	23,0 (26,0)	81,2 (83,1)
2	75,2 (75,8)	100	38,0 (39,0)	86,7 (87,1)
3	73,2 (73,5)	96,9	53,0 (53,0)	84,2 (84,4)
5	83,3 (83,8)	70,8	51,0 (51,0)	76,8 (77,1)
6	63,1 (64,3)	0,0	0,0	0,0
7	60,8 (64,3)	100	32,0 (34,0)	78,0 (80,2)
8	61,7	85,5	77,0	72,6
9	72,0 (72,9)	93,9 (90,9)	35,0 (34,0)	82,2 (81,4)
Average	**71,2 (72,56)**	**91,2 (90,7)**	**44,1 (45,0)**	**80,2 (80,8)**
STD	**7,8 (7,5)**	**10,4 (10,3)**	**17,9 (17,2)**	**4,8 (4,8)**

dissertation. Regarding the performance of the detection, the issue related to the high number of false positives can be observed in almost all the patients.

A detailed analysis of the detection for each patient will then be performed:

Patient 1 - This patient has four of the six FOG episodes completely detected. This results in 72,5% of specificity, 90,9% of sensitivity and 23% of F1-score. The fourth and the last episode missed a single sample at the end. As one can see, many samples before an event are wrongly considered as a FOG. This result is in accordance to what was mentioned in the previous section, about this being a portion of the signal with more tendency to false positives, due to the uncertainties of the labeling process and to a possible phenomenon of initial degradation of gait before the real starting of a FOG - pre-FOG phase.

In addition, episodes that are close, like the third and the fourth episodes and the fifth and sixth, separated by 2,25 and 10 seconds respectively, are considered by the model as a single period of blocking. In fact, one can think that such a small time interval between the end and the beginning of a FOG, such it is 2,25 seconds, may lead to imprecise labeling and wrong decision of the classifier. Figure 5.9 presents the acceleration modulus signal corresponding to the periods of the third and fourth episodes and a few moments before and after their occurrence. The period after the fourth episode is clearly a period of normal gait with easy identified gait cycles. During the periods labeled as FOG the gait degradation can be seen, which is not completely recovered in the time interval between the two episodes. Maybe, a small attempt to resume gait was successfully achieved, but not completely. In any case, since there is no visual information from the DAPHNet dataset one can not be sure of what happened in that moment. Beyond the false detections near the labeled episodes, some punctual noisy detection can also be observed. This noise can easily be removed by the post processing mentioned in Section 3.6. This post processing would increase the metric specificity and the F1-score by 3%, Table 5.3.

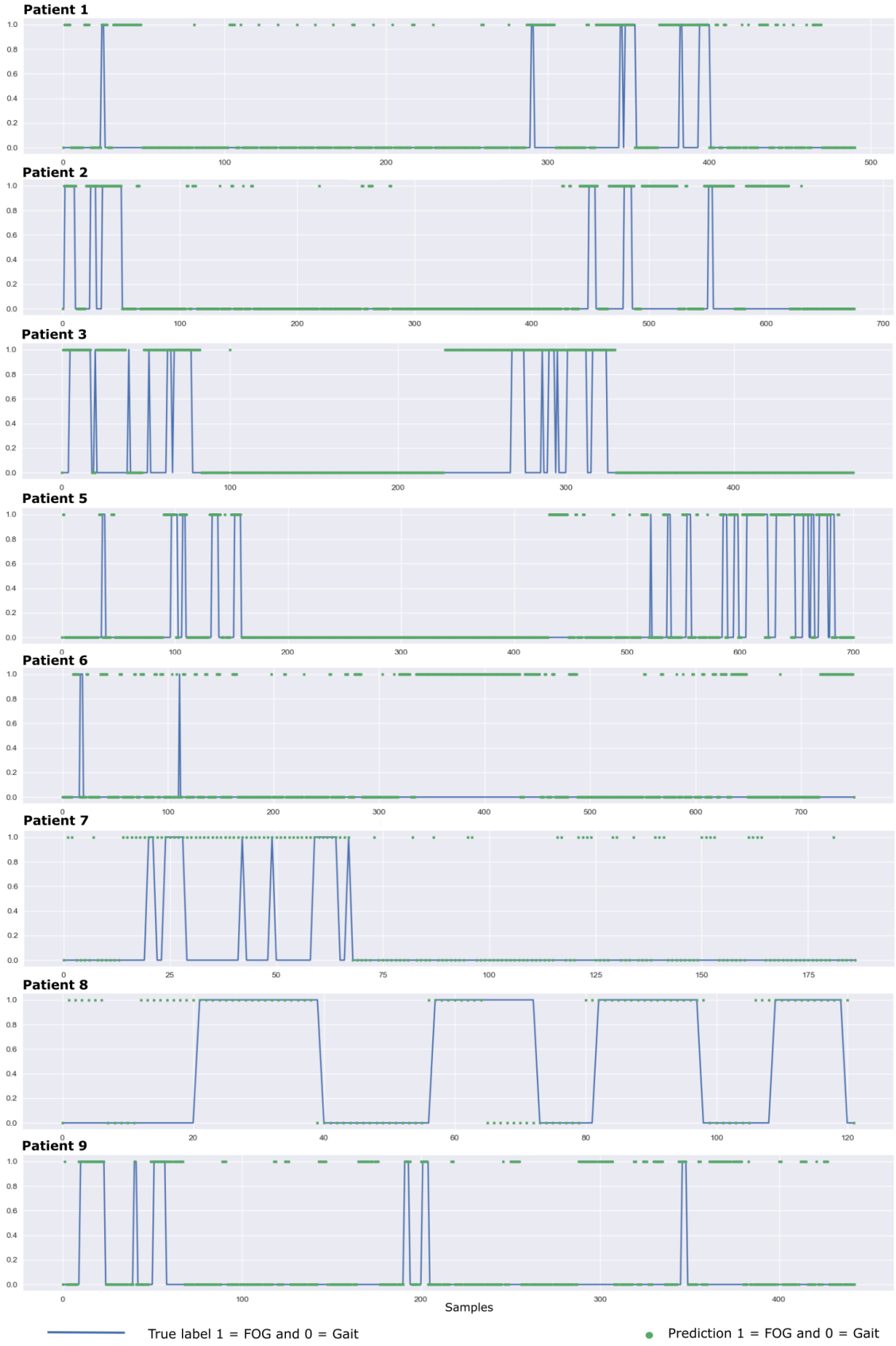

Figure 5.8: Prediction results plotted against the real labels, for the eight considered patients. The small dots represent the prediction and the line represent the real label. The FOG events are identified as 1 and the normal gait as 0.

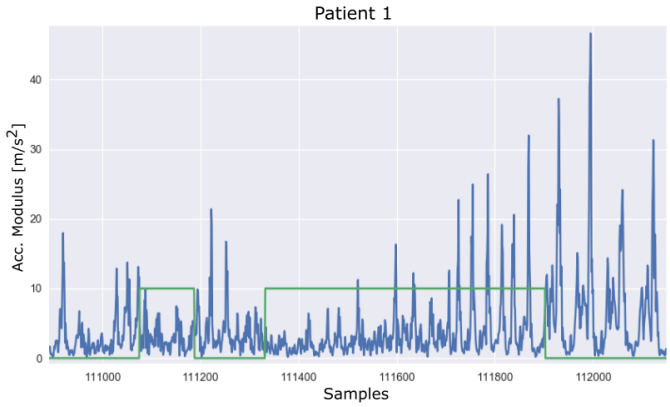

Figure 5.9: Acceleration modulus signal of Patient 1 in the period around the third and the fourth freezing episodes of the test set. The green line represents the label: gait when it is 0 and FOG when it is 10. Between the two episodes, there is not a completely pattern of gait recovery.

Patient 2 - This was the patient that achieved the best classifications performances both in test and train sets. The results in test set were 75,2% of specificity, 100% of sensitivity and 38% of F1-score. All of the six test episodes were completely detected by the model. Similarly to what happened with Patient 1, there are many false positives around the labeled episodes. Episodes 2 and 3, separated by a time interval of 5,19 seconds are detected as a single FOG. The algorithm seems to find two other periods of freezing around the last episode. Although some variations can be observed in the acceleration signal it would be interesting to know its context. Some punctual false detections can also be discarded in the post processing step, increasing the metrics until 1%.

Patient 3 - This patient has a specificity of 73,2%, a sensitivity of 96,9% and a F1-score of 53%. Two out of the twelve episodes were not identified at all. They correspond to very small episodes of FOG (1,11 and 1,50 seconds, respectively). The results of this patient have many false detections between the identified FOGs. It seems that, for this patient, the classifier has difficulties to limit the beginning and the end of an episode. Maybe, the pattern of gait of this patient has different particularities that are not completely distinguished with the features used to get these results. In short, despite the majority of the episodes were detected, the way that the false positives are distributed creating single blocks of freezing is not a desirable result. Moreover, they can not be removed in the post processing step, as they are not punctual false detections.

Patient 5 - The performance metrics of this patient in the test set were: 83,3% of specificity, 70,8% of sensitivity and 51% of F1-score. Among the sixteen episodes of FOG, five were completely detected and, among the remaining, at least 50% of the episode was identified in five of them. In spite of having poor metrics than the patient 3, in the results of this patient a better delimitation of FOG / no FOG periods can be observed. Once again, the wrong detections are concentrated around the period that anticipates each event. Paying attention in the acceleration

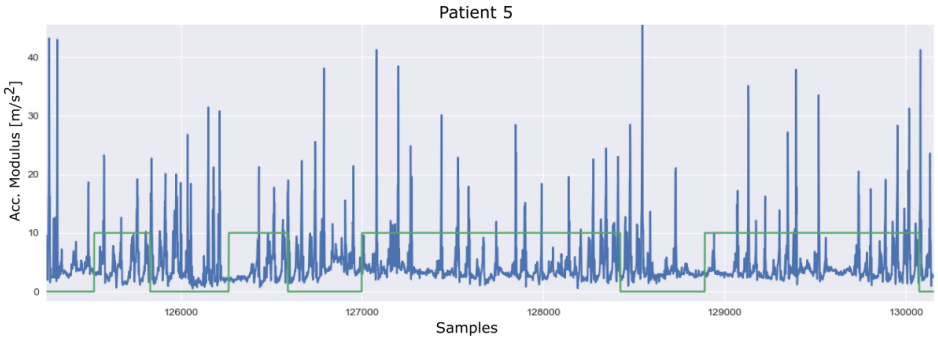

Figure 5.10: Acceleration modulus signal of Patient 5 in the period between the ninth and the twelfth freezing episodes of the test set. The green line represents the label: gait when it is 0 and FOG when it is 10. Before the beginning of the labeled episodes one can see that there is gait degradation that could indicate a period of pre-FOG.

signal, Figure 5.10, it is possible to see variations in the pattern of gait before the starting of the labeled FOG, variations that could represent degradation of gait. Since there is no context about this, one can suppose that it can correspond to the period of pre-FOG or that there are inaccuracies in the labeling of the dataset.

Patient 6 - This was the patient with the poorest results, Table 5.3. They were considered an outlier and thus, were not included in the average performance calculation of the classifier in the test set, as mentioned in Section 5.3.2. Maybe, in this patient, the FOG events were very different among each other and then, due to a relatively low quantity of training samples, the model was not capable of learning FOG characteristics. The test episodes are small, with 2,72 and 1,39 seconds, respectively. None of the them were identified and a large number of false positives exists. In fact, these are very small episodes that some authors would defend that they have no clinical relevance and thus, the fact that they were not identified is not so problematic. The longest continuous period of freezing, wrongly detected, corresponds to a period in the signal when some technical issues, unrelated to the patient gait, happened.

Patient 7 - The results obtained for the Patient 7 in the test set were: 60,8% of specificity, 100% of sensitivity and 32% of F1-score. All the episodes were completely detected, but the prediction joined them into a single long period of freezing, enable to identify the correct start and end of the episodes. A period of freezing is identified instead of individual events of freezing. The noisy detections can be eliminated in the post-processing step, increasing the F1-score in 2% and the specificity in almost 4%.

Patient 8 - This patient achieved a specificity of 61,7%, a sensitivity of 85,5% and a F1-score of 77%. Two of the four episodes were completely identified with a few false positives around the true labels. The other two FOG were almost completely identified. Analyzing Figure 5.8, one can

say that the model was able to identify the events of freezing in a very good way, with a relatively low number of false positives and without noise. Consequently, the post-processing step here has no effect, as one can see in Table 5.3.

Patient 9 - Lastly, Patient 9 had 72,0% of specificity, 93,9% of sensitivity and a F1-score of 35%. Among the six test episodes, four of them were completely identified and the other two are partially detected - more than 50%. Observing the figure, one can see that most of the episodes are very well detected and limited, with just a few false positives around each episode. Some false positives also exist in the middle of the gait periods. In this specific case, the application of the post-processing step eliminates also some true positive detections, decreasing the value of sensitivity in 3%. So, in this case, this step is not favorable. In average, in the set of patients tested, this step increases the performance results by approximately 1%.

Another interesting result is related with the best SVM parameters found for each patient. For 50% of them, the same parameters were chosen : $C = 1$ and $\gamma = 0,1$. Having the possibility of doing more tests with different patients, if the tendency continues to be verified, the exhaustive step of tuning the parameters can be ignored, by defining the classifier with the most common parameters.

Concerning the high number of false positives, it is important to notice that a few attempts have been made to reduce them. In the beginning of the results acquisition, they were extracted considering the same weight to both classes, instead of giving more weight to the FOG class, which is under-represented. In other words, the SVM was trained considering that both gait and FOG samples have the same relevance. This led to a lower number of false positives, but also, to a lower number of true positives, since sometimes, in an entire episode of FOG, only a sample was correctly detected. Consequently the model sensitivity decreased. Other attempts also resulted in this issue - removing the number of false positives led to a decrease in the number of true positives. So a compromise needs to be defined, taking into account what is the wrong detection that costs more: a gait sample being identified as a FOG sample, or a FOG sample being identified as gait. Due to the nature of FOG and to the burden of these events for the patients, a FOG sample being identified as gait is then more problematic. Moreover, it is interesting to refer that in a related study, Mazilu *et al.* [8] have tested a wearable system for FOG detection in a set of 5 patients and they have concluded that sometimes what is counted as false positives is perceived as helpful by users, since sometimes the alert triggered by a false positive detection helped the patients to avoid a potential FOG. This could show that maybe a slight degradation of the gait, that it is not yet considered as a blocking moment (pre-FOG), is detected by the algorithm. These results and conclusions clearly indicate that more experiments, with more patients, need to be performed in order to gather more information.

5.4 Pre-FOG Analysis Results

This section reports the results of the statistical analysis performed in the pre-FOG periods. As described in Chapter 4, a statistical analysis was performed for each feature extracted in each patient. The goal was to evaluate if there are significant differences between its values during gait and during pre-FOG and consequently, conclude about the existence or not of this period.

Figure 5.11 presents the obtained results, identifying for each patient the features that have shown significant differences.

	HHF				HHA				SHHF				SHHA				DiffH				SM				
Patient	MOD	HFA	VA	HLA	MOD	HFA	VA	HLA	MOD	HFA	HLA	VA	MOD	HFA	HLA	VA	MOD	HFA	HLA	VA	MOD	HFA	VA	HLA	Total
1		X	X				X			X						X					X	X	X		8
2		X			X	X	X		X				X			X							X		8
3																									0
5		X	X	X	X	X	X	X	X	X	X		X	X		X	X	X			X	X	X	X	19
6						X	X						X	X		X					X	X		X	8
7	X	X	X			X	X					X	X	X		X	X				X	X	X	X	14
8		X	X			X						X		X								X	X		7
9	X	X	X	X		X	X			X		X	X	X		X	X		X		X	X	X		16
Total	2	6	5	2	2	6	6	1	2	3	1	3	5	5	0	6	3	1	1	0	5	6	6	3	

Figure 5.11: Set of statistical tests performed. The ones marked with a 'X' are the ones that show statistical differences between normal gait and FOG.

Analyzing the results, the first aspect to notice is that none of the features extracted from Patient 3 has shown to have significant difference between the two periods and thus, there is no potential to distinguish between them. From these results, one can suppose that, for the concerned patient, the transition from normal gait to freezing is very quick and thus, a period where the gait gradually degrades until the FOG is reached does not exist.

On the contrary, Patient 5 is the one with a higher number of significant features. From the twenty four extracted features, nineteen of them seem to have a potential to predict FOG before its beginning. The similarity measures and the amplitude of the higher harmonic in the four signals are completely represented in the set of features with significant differences. The results obtained for this patient are in agreement with those found in the previous section. Going back to the detection results, one can see that the false positives were concentrated in the periods before the beginning of the labeled FOG events, which was already suggesting the existence of a pre-FOG period. Moreover, the analysis performed on the acceleration signals before freezing events also suggested that this patient has variations on the walking pattern before the blocking state.

The remaining patients also have a set of features with potential to detect a pre-FOG.

Figures 5.12 and 5.13 show the paired t-test results for an illustrative set of features, for Patient 5 and 3. Since the statistical analysis comprised in total 24×8 paired t-tests, only illustrative examples of these tests, which leaded to the results presented in the figures, are shown. The graphic on the left presents the values of each pair gait - pre-FOG considered in the statistical test. On the right, the mean and standard deviation are reported.

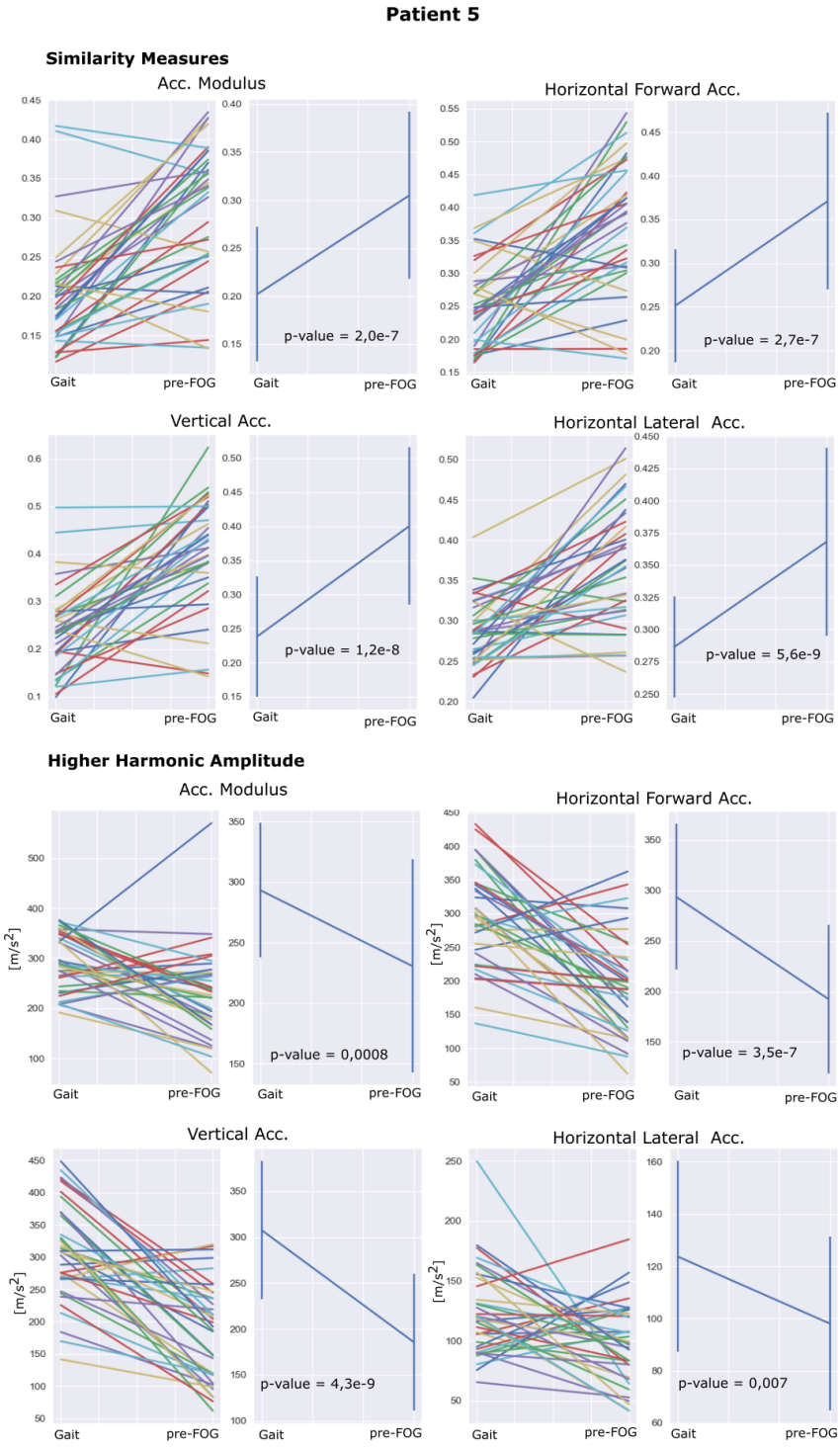

Figure 5.12: Results of the statistical analysis performed with the Patient 5 features. Only the similarity measures and the higher harmonic amplitude computed with the four acceleration signals are presented, as illustrative statistical tests. For each feature, the graphic on the left represents the feature value for each pair gait - pre-FOG and on the right the mean and standard deviation. The horizontal axis of the graph has no physical meaning, since it represents two categorical values (gait and pre-FOG)

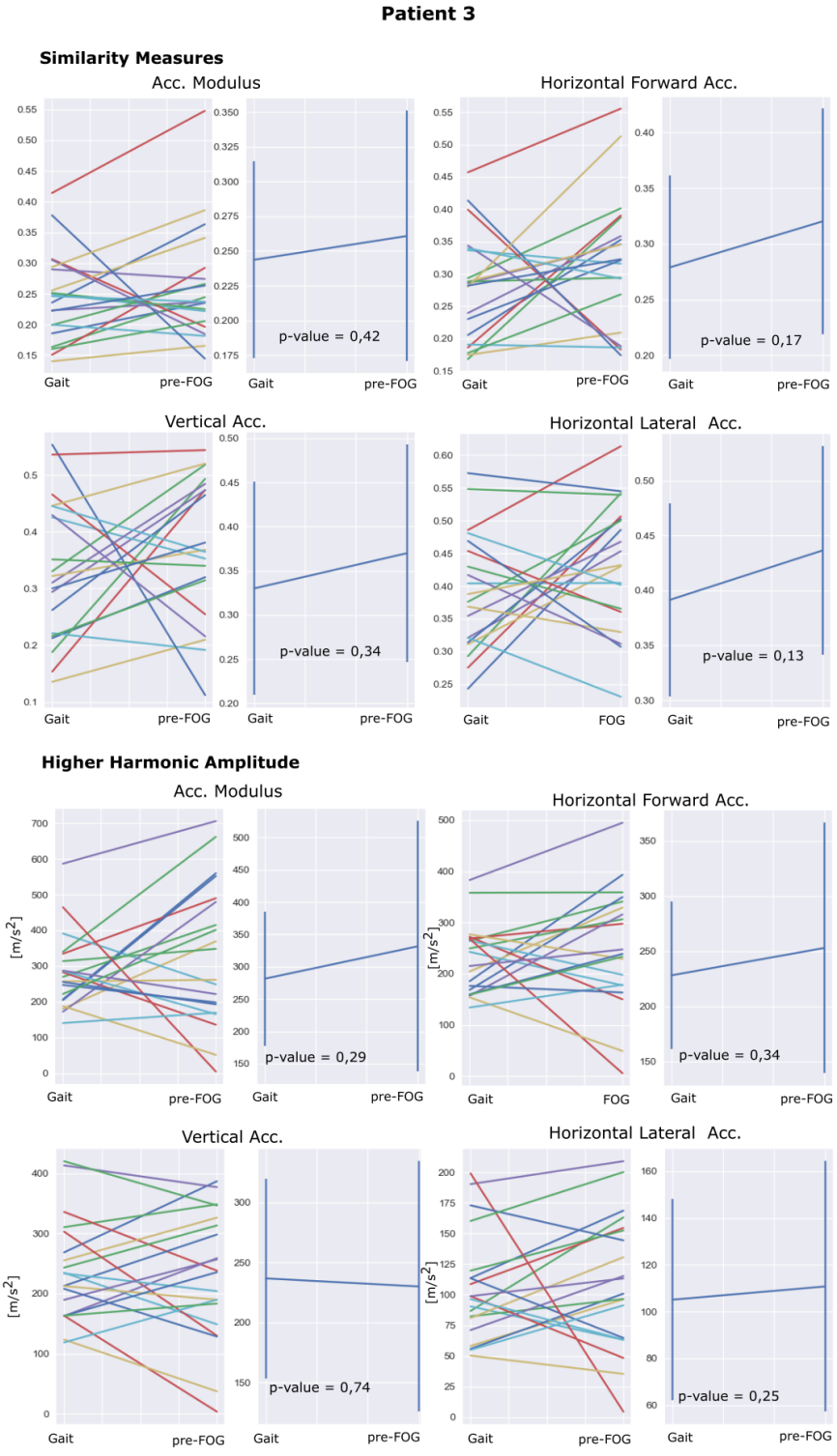

Figure 5.13: Results of the statistical analysis performed with the Patient 3 features. Only the similarity measures and the higher harmonic amplitude computed with the four acceleration signals are presented, as illustrative statistical tests. For each feature, the graphic on the left represents the feature value for each pair gait - pre-FOG and on the right the mean and standard deviation. The horizontal axis of the graph has no physical meaning, since it represents two categorical values (gait and pre-FOG)

The features from Patient 5 have significant differences between gait and pre-FOG, since the p-value is smaller than 0,05. As one can see, from the results of the similarity measures, in the set of pairs considered, there is a tendency to have smaller values of the DTW during gait, than during the period of pre-FOG (here considered as 3 windows before the beginning of FOG). These results make sense, taking into account that low values of DTW mean more similarity between the sequences, that is what happens when comparing the created template with the normal gait windows. If the pattern of gait starts to vary before the occurrence of a freezing event, then the similarity with the template will decrease. This tendency is visible for the four similarity features, as one can see from the figures on the right. The amplitudes of the higher harmonics also show a consistent behavior, with a decrease in their values in the windows corresponding to pre-FOG.

It is important to clarify that, in the case of the performed tests, the STD bars only represent how much scatter each group (gait and pre-FOG) has, but it says nothing about the significance of the difference between the two groups. It happens because a dependent t-test was performed, where repeated measures (within-subjects) were computed and thus, what matters is the consistency of the differences. Consequently, STD bars tells nothing about the significance of the test [107].

Concerning the results obtained for Patient 3, from the graphics on the left, it can be seen that there is a greater variety in the behavior of the features, leading to no significant differences in the values during gait and pre-FOG: the p-values are all above 0,05.

Focusing now on individual features, the similarity measures are the ones more frequently identified as having significant differences between pre-FOG and normal gait, followed by the amplitude of the harmonics. On the other hand, the difference between the frequency of the first and second harmonics is poorly represented in these results.

5.5 Computational Cost and Real Time Implementation

The similarity measures proposed in this dissertation have, as mentioned in Section 3.4.1.1, a relatively high computational complexity. This complexity decreases with the FastDTW algorithm implementation, but yet, evaluating the computational cost of their extraction can be important, since one of the goals would be the algorithm integration in a real time application.

Then, the computation cost of extracting of the proposed features is compared with the computation cost of extracting the Mazilu's. This comparison was performed by calculating the time needed to complete the both sets of features extraction. Moreover, the time needed to do signal segmentation based on the detection of gait cycles and based on fixed time windows was also compared. It is important to notice that this time depends on the processor capabilities and on how busy it is with other tasks. In this way, the analysis of the relative values is more important than the absolute values of time.

The results indicate that the signal segmentation based on the detection of the gait cycles is, in average, 1,8 times slower than the traditional methods for signal segmentation. Additionally, the

extraction of the features proposed in this work is, on average, 1,7 times slower than the extraction of those proposed by Mazilu.

Hence, the results point out a higher computational burden of the proposed approach. This is not necessarily a bad result, since in both approaches, optimization steps should be necessary to implement them in a real-time application.

Even though a real-time implementation of the algorithm was not performed in this work, it is important to notice that except for the post-processing step, all the proposed approaches are able to be applied in a real-time context. Nevertheless, some adjustments may be needed. In the pre-processing step, the performed operations (detrending and filtering) are now applied with the totality of the signal. In a real-time implementation, they should be applied individually in each window and so, other types of filters may be needed. The signal segmentation step, that is now based in the detection of peaks that may correspond to the beginning of the stance phase of the gait, may also need small adjustments. The subsequent windows division that is performed to deal with those that do not represent individual gait cycles, could be done simultaneously with the peak detection. In other words, in real-time, if a peak is not detected in a period corresponding to two times the mean of a gait cycle, a window is created, instead of waiting for the detection of the next peak. The implementation of these adjustments is part of the future work of this dissertation.

Finally, concerning the latency of the algorithm, a few considerations can also be done. The similarity measures are computed from a single gait cycle window and the frequency-based features are extracted from a time interval of three windows. Excepting for the initial moments of acquisition, the delay associated to the algorithm is of one gait cycle, that depending of the patient could represent a period of time from about 1 to 2 seconds.

Chapter 6

Conclusions and Future Work

During the last few years, many studies explored the monitoring of the motor PD symptoms, such as FOG, through the use of inertial sensors systems. The studies aim to provide the physicians with efficient tools for both monitoring and management of freezing episodes. Most of these tools try to detect FOG, or even predict its occurrence, allowing the application of auditory stimulation that helps the patients to resume gait. Despite the number of studies performed in this field, many explorations and improvements can be done to increase the efficiency of FOG detection.

This dissertation presented a supervised machine learning methodology to detect FOG using inertial data from accelerometers placed on the lower limb of the patient, namely on an ankle. It is a patient-dependent approach and a new segmentation methodology, and also a new set of features, were proposed. The algorithm was tested with inertial data of 8 PD patients, belonging to a public dataset - DAPHNet dataset. A previous understanding of the FOG phenomenon, achieved by research and analysis of the acceleration signals presented in the dataset, was an important phase of the dissertation work and has shown to be crucial in the design of the proposed methodology.

Therefore, the signal segmentation is performed based on the detection of the patient gait cycles and then, the extraction of similarity measures from gait patterns and of frequency-based features is performed. Until the writing of this report, the similarity measures, computed from template matching approaches, were only used one time in the context of FOG detection, in a methodology quite different from the one proposed in this work. Hence, no related work has used this kind of measures as features to the input of classifiers. Furthermore, the proposed approach has been compared to other studies presented in the literature that achieved significant results.

With the proposed patient-dependent methodology a sensitivity of 85,1% and a specificity of 82,0% were achieved along with a F1-score of 58,9%. Taking into account that most of the studies only report the results in terms of sensitivity and specificity, comparing these values, one can conclude that they either present similar results or outperform previous studies. Consequently, the presented results suggest that the proposed methodology can be used to detect FOG events in PD patients. They are a preliminary evidence for the feasibility of using similarity measures and template based approaches to distinguish between FOG and normal gait, along with other frequency-based features. These values represent an average of the performance of the classifier

in the 8 tested patients. The individual results show that the classification algorithm performed well for some patients, but had not so satisfactory results for others. In this way, future work should be done to better validate and improve the results.

The main limitation of this study is the reduced amount of data to be used in a patient-dependent approach and the small number of subjects involved. A phenomenon with such a patient dependent nature should be studied in a larger population of subjects in order to draw more reliable conclusions about the results. One of the initial goals of this dissertation was to be part of a new dataset acquisition, where one could better validate the reported results. Moreover, the acquisition of more data would be useful not only to validate the results but also to a better understanding of the freezing events. Observing how they occur, the context that led to a freezing event and additionally, having this information annotated in the dataset, would be extremely important and could open the door to the exploration of new ways of improving the presented results. Ultimately, this acquisition was not performed because of the bureaucratic issues associated, that originated a long wait for the ethical approvals from the Ethical Committee of Hospital S.Joao. Thus, this project was performed using only acceleration data from the DAPHNet dataset.

The acquisition of new data would also include new elements for further exploration. Signals from accelerometers and gyroscopes would be recorded from three different locations: both ankles, trunk and head. The periods of freezing and their context and also the periods of voluntary absence of movement would be labeled. Notice that a few recent works are already exploring the contribution of gyroscope data both for FOG and pre-FOG detection, reporting relatively good results. Hence, creating a dataset with a more complete set of information, can contribute to the improvement of the FOG detection methodology.

The algorithm needs some optimization in order to be implemented in wearable real-time applications. It was designed to use acceleration data acquired from the sensor placed on the ankle. That is, the segmentation of the signal into gait cycles is performed based on the gait pattern of the ankle acceleration. In order to be applied in data from different locations, a few adjustments in the gait cycle detection algorithm are needed.

Concerning the analysis performed on the pre-FOG period, one can conclude that some patients manifest degradation of gait prior to FOG that could be detected through the proposed features. This finding is for some patients reinforced by the existence of many false positives in the period before a FOG, as observed in the detection results.

As a future work task, it would be interesting to analyze the performance of a classifier in the detection of these regions. Here the validation of the results would be more difficult to do since one cannot label the pre-FOG periods. Different features associated to gait parameters and computed from the combination of the acceleration and gyroscope signal could also be explored as well as the most appropriate duration of pre-FOG for each patient.

Appendix A

Study Documentation

A.1 Research Submission

Comissão de Ética Centro Hospitalar São João / Faculdade de Medicina da Universidade do Porto n.º 36 / 18

SÃO JOÃO

U.PORTO FMUP FACULDADE DE MEDICINA UNIVERSIDADE DO PORTO

Questionário para submissão de Investigação

Exmo. Sr. Presidente da Comissão de Ética do Centro Hospitalar de São João/ Faculdade de Medicina da Universidade do Porto,

Pretendendo realizar a investigação infracitada, solicito a V. Exa., na qualidade de Investigador, a sua apreciação e a elaboração do respetivo parecer. Para o efeito, anexo toda a documentação requerida.

IDENTIFICAÇÃO DO ESTUDO

Título da investigação: FOGSensor4PD - Real-time detection of FOG episodes in patients with Parkinson's Disease

Nome do investigador: Vânia Margarida Cardoso Guimarães

Endereço eletrónico: vania.guimaraes@fraunhofer.pt **Contacto telefónico:** 220430348

Caracterização da investigação:

☐ Estudo retrospetivo ☐ Estudo observacional ☐ Estudo prospetivo

☐ Inquérito ☒ Outro. Qual? Aquisição de dados - aprendizagem supervisionada

Tipo de investigação:

☐ Com intervenção ☒ Sem intervenção

Formação do investigador em boas práticas clínicas *(GCP)*: ☐ Sim ☒ Não

Promotor *(se aplicável)*: Associação Fraunhofer Portugal Research

Nome do orientador de dissertação/tese *(se aplicável)*: Vânia Margarida Cardoso Guimarães

Endereço eletrónico: vania.guimaraes@fraunhofer.pt

Local/locais onde se realiza a investigação: Centro Hospitalar de São João

Data prevista para início: 01 / 02 / 2018 Data prevista para o término: 31 / 07 / 2018

PROTOCOLO DO ESTUDO

Síntese dos objetivos:

Desenvolvimento de um sistema wearable, portátil, para deteção e/ou previsão de episódios de congelamento da marcha (ou episódios FOG, do inglês 'Freezing of Gait') em pessoas com a doença de Parkinson. Para tal, serão utilizados sensores inerciais colocados em partes estratégicas do corpo, por forma a capacitar a avaliação do movimento na doença de Parkinson. Nesta primeira fase, o estudo pretende realizar aquisição de dados de marcha e episódios de FOG em doentes com Parkinson, em condições laboratoriais, com vista ao desenvolvimento de algoritmos capazes de detetar e/ou prever este tipo de eventos.

Fundamentação ética (ganhos em conhecimento/inovação; ponderação benefícios/riscos):

A ocorrência de episódios de FOG é comum em estágios mais avançados da doença de Parkinson. Estes episódios são caracterizados por uma pausa imprevisível e transitória na marcha, que ocorre quando o doente começa a caminhar, ou enquanto caminha, particularmente ao mudar de direção ou ao passar em locais estreitos. A ocorrência destes episódios está associada a restrição de mobilidade e aumento do risco de queda, o que leva a uma perda progressiva da independência e qualidade de vida dos doentes. A deteção automática destes sintomas permite um acompanhamento objetivo da progressão da doença, e uma melhor gestão da medicação. Para além disso, a capacidade de detetar e, mais ainda, prever a ocorrência deste tipo de episódios permite uma mais eficaz integração com sistemas de auxílio à marcha, por forma a que sejam capazes de atuar apenas quando necessário e prevenir ou encurtar a duração dos episódios de FOG. Estas funcionalidades, além de inovadoras, constituem um importante avanço tecnológico no autocuidado e gestão da doença de Parkinson, permitindo prolongar a independência e autonomia do doente, bem como melhorar a sua qualidade de vida. Os sensores que pretendemos utilizar são externos, não invasivos, não constituindo nenhum risco nem causando qualquer incómodo aos participantes. Para além disso, as atividades que propomos realizar no âmbito do projeto constituem atividades normais do dia-a-dia dos pacientes, ✚

1/3

CES-IM007-0

CONFIDENCIALIDADE

De que forma é garantida a anonimização dos dados recolhidos de toda a informação?

Os pacientes serão identificados por um código, de modo a que os dados recolhidos jamais possam ser associados por terceiros à pessoa a que se referem. Por força de contrato com a Fraunhofer, os investigadores são obrigados a manter sigilo e confidencialidade dos dados

O investigador necessita ter acesso a dados do processo clínico? ☒ Sim ☐ Não

Está previsto o registo de imagem ou som dos participantes? ☒ Sim ☐ Não

Se sim, está prevista a destruição deste registo após o sua utilização? ☒ Sim ☐ Não

CONSENTI MENTO

O estudo implica recrutamento de:

Doentes: ☒ Sim ☐ Não Voluntários saudáveis: ☐ Sim ☒ Não

Menores de 18 anos: ☐ Sim ☒ Não

Outras pessoas sem capacidade do exercício de autonomia: ☐ Sim ☒ Não

A investigação prevê a obtenção de Consentimento Informado: ☒ Sim ☐ Não

Se não, referir qual o fundamento para a isenção:

Existe informação escrita aos participantes: ☒ Sim ☐ Não

PROPRIEDADE DOS DADOS

A investigação e os seus resultados são propriedade intelectual de:

☐ Investigador ☒ Promotor ☐ Ambos ☐ Serviço onde é realizado

☐ Não aplicável Outro: _____

BENEFÍCIOS, RISCOS E CONTRAPARTIDAS PARA OS PARTICIPANTES

Benefícios previsíveis:

Os participantes não terão nenhum benefício imediato, contudo, estarão a contribuir para o desenvolvimento de tecnologia que permita melhorar a gestão e autocuidado na doença de Parkinson, que possa também contribuir para a manutenção da mobilidade e independência do doente.

Riscos/incómodos previsíveis:

Serão utilizados sensores externos não invasivos, já comercializados, não constituindo nenhum risco nem incómodo aos participantes. Por razões de segurança, haverá um profissional de saúde dedicado ao acompanhamento dos doentes durante os testes, os quais incluem apenas atividades normais do dia-a-dia. Os testes serão realizados no mesmo dia da consulta já agendada, pelo que os

São dadas contrapartidas aos participantes:

- ***pela participação*** ☐ Sim ☒ Não ☐ Não aplicável
- ***pelas deslocações*** ☐ Sim ☐ Não ☒ Não aplicável
- ***pelas faltas ao emprego*** ☐ Sim ☐ Não ☒ Não aplicável
- ***por outras perdas e danos*** ☐ Sim ☐ Não ☒ Não aplicável

CUSTOS / PLANO FINANCEIRO

Os custos da investigação são suportados por:

☐ Investigador ☒ Promotor ☐ Serviço onde é realizado

☐ Não aplicável Outro: _____

Existe protocolo financeiro? ☐ Sim ☒ Não

LISTA DE DOCUMENTOS ANEXOS

- ☒ Pedido de autorização ao Presidente do Conselho de Administração do Centro Hospitalar de São João *(se aplicável)*
- ☐ Pedido de autorização à Diretora da Faculdade de Medicina da Universidade do Porto *(se aplicável)*
- ☒ Protocolo do estudo
- ☒ Declaração do Diretor de Serviço onde decorre o estudo
  (sendo um estudo na área de enfermagem deve anexar também a concordância da chefia de enfermagem)
- ☒ Profissional de ligação
- ☒ Informação dos orientadores
- ☒ Informação ao participante
- ☒ Modelo de consentimento
- ☒ Instrumentos a utilizar *(inquéritos, questionários,escalas, p.ex.)*: Questionário e escalas incluídas no questionário
- ☒ Curriculum Vitae abreviado *(máx. 3 páginas)*
- ☐ Protocolo financeiro
- ☒ Outros:

Ficha Clínica Avaliável (FCA), Declaração do Orientador da Instituição de Investigação, Equipa de Investigação

COMPROMISSO DE HONRA E DECLARAÇÃO DE INTERESSES

Declaro por minha honra que as informações prestadas neste questionário são verdadeiras. Mais declaro que, durante o estudo, serão respeitadas as recomendações constantes da Declaração de Helsínquia (1960 e respetivas emendas), e da Organização Mundial da Saúde, Convenção de Oviedo e das "Boas Práticas Clínicas" (GCP/ICH) no que se refere à experimentação que envolve seres humanos. Aceito, também, a recomendação da CES de que o recrutamento para este estudo se fará junto de doentes que não tenham participado em outro estudo, nos últimos três meses. Comprometo-me a entregar à CES o relatório final da investigação, assim que concluído.

Porto, 24 *de* Janeiro *de* 2018

Nome legível: Vânia Margarida Cardoso Guimarães Vânia Guimarães

assinatura

Parecer da Comissão de Ética do Centro Hospitalar de São João/ FMUP Emitido na reunião plenária da CE de 16 / 02 / 18

A Comissão de Ética para a Saúde APROVA por unanimidade o parecer do Relator, pelo que nada tem a opor à realização deste projecto de investigação.

Prof. Doutor Filipe Almeida
Presidente da Comissão de Ética

A.2 Research Approvals

Unidade de Investigação

Tomei conhecimento. Nada a opor.

20 de Abril de 2018

A Coordenadora da Unidade de Investigação

(Prof.ª Doutora Ana Azevedo)

Aprovado. Ao CA.

(Prof.ª Doutora Ana Azevedo)

n.º 36 / 18

SÃO JOÃO

PEDIDO DE AUTORIZAÇÃO

Realização de Investigação

Exmo. Senhor Presidente do Conselho de Administração
do Centro Hospitalar de São João

AUTORIZADO

CONSELHO DE ADMINISTRAÇÃO REUNIÃO DE 26 ABR 2018

Presidente do Conselho de Administração

(Dr. António Oliveira e Silva)

Diretor Clínico | Enfermeira Diretora | Vogal Executivo | Vogal Executivo

(Prof. Dr. João Artur Costa) | (Enf.ª Filomena Cardoso) | (Dr. Luís Pinto Correia) | (Dr. Rui Sá G. Mateus)

Nome do Investigador Principal:

Vânia Margarida Cardoso Guimarães

Título da Investigação:

FOGSensor4PD - Real-time detection of FOG episodes in patients with Parkinson's Disease

Pretendendo realizar no(s) Serviço(s) de:

Medicina Física e de Reabilitação

a investigação em epígrafe, solicito a V. Exa., na qualidade de Investigador/Promotor, autorização para a sua efetivação.

Para o efeito, anexo toda a documentação referida no dossier da Comissão de Ética do Centro Hospitalar de São João/ Faculdade de Medicina da Universidade do Porto respeitante à investigação, à qual enderecei pedido de apreciação e parecer.

Com os melhores cumprimentos. *O Investigador/ Promotor*

Porto, 24 *de* Janeiro *de* 2018 . Vânia Guimarães

assinatura

CES-IM005-0

SÃO JOÃO
HOSPITAL

Parecer da Comissão de Ética para a Saúde do
Centro Hospitalar de São João / Faculdade de Medicina da Universidade do Porto

Título do Projecto: FOGSensor4PD - Real-time detection of FOG episodes in patients with Parkinson's Disease.

Nome da Investigadora Principal: Eng.ª Vânia Margarida Cardoso Guimarães, investigadora da Fraunhofer Portugal.

Onde decorre o Estudo: No Serviço de Medicina Física e Reabilitação do CHSJ. Dispõe de autorização do Dr. Fernando Parada, e terá como profissional de ligação a Dra. Maria José Festas.

Objectivos do Estudo: Estudo promovido pela Associação Fraunhofer Portugal Research cujo principal objectivo consiste no desenvolvimento de um sistema wearable, portátil, para detecção e/ou previsão de episódios de "congelamento da marcha" em pessoas com a doença de Parkinson.

Para o efeito, numa primeira fase prevê-se o recrutamento de 10 a 20 participantes, a quem serão colocados sensores do movimento em zonas específicas do corpo para monitorizar episódios de congelamento da marcha, com vista ao desenvolvimento de algoritmos para a detecção destes eventos, cuja monitorização será realizada com ajuda de uma aplicação de smartphone.

Concepção e Pertinência do estudo: A pertinência do estudo assenta nos benefícios para o autocuidado e gestão da doença de Parkinson através da prevenção de episódios de que "congelamento da marcha". Não será solicitado aos doentes que realizem qualquer atividade para além da sua atividade de rotina diária.

Benefício/risco: Não existem riscos, benefícios ou incómodos diretamente relacionados com o estudo, que se espera que venha a contribuir para dispor no futuro de instrumentos de auxílio à marcha que melhorem a qualidade de vida destes doentes. Não estão previstas deslocações dos doentes, especificamente para a participação estudo.

Confidencialidade dos dados: A cada paciente será atribuído um código de participação, garantindo o seu anonimato. Está previsto o acesso aos dados clínicos através do elo de ligação. Está também prevista a obtenção de registos de imagem, que serão destruídos no final do estudo.

Respeito pela liberdade e autonomia do sujeito de ensaio: Está prevista a obtenção de consentimento informado, o qual é acompanhado de uma informação escrita para os participantes que é esclarecedora sobre a natureza do estudo e contempla as questões éticas relevantes.

***Curriculum* da investigadora**: Adequado à investigação.

Data previsível da conclusão do estudo: Julho de 2018

Conclusão: Proponho um parecer favorável à realização deste projecto de investigação.

Porto, 16 de Fevereiro de 2018

O Relator da CES, Prof. Doutor Manuel Pestana

Bibliography

[1] S. C. Mukhopadhyay, "Wearable sensors for human activity monitoring: A review," *IEEE Sensors Journal*, vol. 15, no. 3, pp. 1321–1330, 2014.

[2] S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers, "A review of wearable sensors and systems with application in rehabilitation," *Journal of NeuroEngineering and Rehabilitation*, vol. 9, no. 1, p. 21, 2012. [Online]. Available: http://jneuroengrehab.biomedcentral.com/articles/10.1186/1743-0003-9-21

[3] P. Lorenzi, R. Rao, G. Romano, A. Kita, M. Serpa, F. Filesi, and F. Irrera, "Smart sensors for the recognition of specific human motion disorders in Parkinson's disease," pp. 131–136, 2015.

[4] E. Rovini, C. Maremmani, and F. Cavallo, "How wearable sensors can support parkinson's disease diagnosis and treatment: A systematic review," *Frontiers in Neuroscience*, vol. 11, no. OCT, 2017.

[5] L. Palmerini, L. Rocchi, S. Mazilu, E. Gazit, J. M. Hausdorff, and L. Chiari, "Identification of characteristic motor patterns preceding freezing of gait in Parkinson's disease using wearable sensors," *Frontiers in Neurology*, vol. 8, no. AUG, pp. 1–12, 2017.

[6] A. L. Silva de Lima, L. J. W. Evers, T. Hahn, L. Bataille, J. L. Hamilton, M. A. Little, Y. Okuma, B. R. Bloem, and M. J. Faber, "Freezing of gait and fall detection in Parkinson's disease using wearable sensors: a systematic review," *Journal of Neurology*, vol. 264, no. 8, pp. 1642–1654, 2017. [Online]. Available: http://link.springer.com/10.1007/s00415-017-8424-0

[7] W.-H. Wang, P.-C. Chung, Y.-L. Hsu, M.-C. Pai, and C.-W. Lin, "Inertial-Sensor-Based Balance Analysis System for Patients with Alzheimer's Disease," *Conference on Technologies and Applications of Artificial Intelligence*, pp. 128–133, 2013. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6783855

[8] S. Mazilu, U. Blanke, M. Hardegger, G. Tröster, E. Gazit, and J. M. Hausdorff, "GaitAssist: A Daily-Life Support and Training System for Parkinson's Disease Patients with Freezing of Gait," *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*, pp. 2531—-2540, 2014.

[9] J. Jankovic, "Parkinson's disease: Clinical features and diagnosis," *Journal of Neurology, Neurosurgery and Psychiatry*, vol. 79, no. 4, pp. 368–376, 2008.

[10] Y. Okuma, "Freezing of gait in Parkinson's disease," *Journal of Neurology*, vol. 253, no. SUPPL. 7, pp. 27–32, 2006.

[11] P. Arias and J. Cudeiro, "Effect of Rhythmic Auditory Stimulation on Gait in Parkinsonian Patients with and without Freezing of Gait," *PLoS ONE*, vol. 5, no. 3, 2010.

[12] A. Nieuwboer, K. Baker, A.-M. Willems, D. Jones, J. Spildooren, I. Lim, G. Kwakkel, E. Van Wegen, and L. Rochester, "The Short-Term Effects of Different Cueing Modalities on Turn Speed in People with Parkinson's Disease," *Neurorehabilitation and Neural Repair*, vol. 23, no. 8, pp. 831–836, 2009. [Online]. Available: http://journals.sagepub.com/doi/10.1177/1545968309337136

[13] C. G. Goetz, B. C. Tilley, and S. R. Shaftman et al, "Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results," *Movement Disorders*, vol. 23, no. 15, pp. 2129–2170, 2008.

[14] N. Giladi, H. Shabtai, E. S. Simon, S. Biran, J. Tal, and A. D. Korczyn, "Construction of freezing of gait questionnaire for patients with Parkinsonism," *Parkinsonism & related disorders*, vol. 6, pp. 165—-170, 2000.

[15] A. Samà, D. Rodríguez-Martín, C. Pérez-López, A. Català, S. Alcaine, B. Mestre, A. Prats, M. C. Crespo, and À. Bayés, "Determining the optimal features in freezing of gait detection through a single waist accelerometer in home environments," *Pattern Recognition Letters*, vol. 0, pp. 1–9, 2017.

[16] M. Bachlin, M. Plotnik, D. Roggen, I. Maidan, J. Hausdorff, N. Giladi, and G. Troster, "Wearable Assistant for Parkinson's Disease Patients With the Freezing of Gait Symptom," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 436–446, 2010. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs{_}all.jsp?arnumber=5325884{%}5Cnhttp://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5325884

[17] D. Rodríguez-Martín, C. Pérez-López, A. Samà, A. Català, J. Moreno Arostegui, J. Cabestany, B. Mestre, S. Alcaine, A. Prats, M. Cruz Crespo, and À. Bayés, "A Waist-Worn Inertial Measurement Unit for Long-Term Monitoring of Parkinson's Disease Patients," *Sensors*, vol. 17, no. 4, p. 827, 2017. [Online]. Available: http://www.mdpi.com/1424-8220/17/4/827

[18] D. Rodríguez-Martín, A. Samà, C. Pérez-López, A. Català, J. M. Arostegui, J. Cabestany, À. Bayés, S. Alcaine, B. Mestre, A. Prats, M. C. Crespo, T. J. Counihan, P. Browne,

L. R. Quinlan, G. Laighin, D. Sweeney, H. Lewy, J. Azuri, G. Vainstein, R. Annicchiarico, A. Costa, and A. Rodríguez-Molinero, "Home detection of freezing of gait using Support Vector Machines through a single waist-worn triaxial accelerometer," *PLoS ONE*, vol. 12, 2017.

[19] S. Mazilu, A. Calatroni, E. Gazit, D. Roggen, J. M. Hausdorff, and G. Tröster, "Feature Learning for Detection and Prediction of Freezing of Gait in Parkinson's Disease," in *International Workshop on Machine Learning and Data Mining in Pattern Recognition*, 2013, pp. 144–158.

[20] F. Irrera, A. Kita, R. Rao, and A. Suppa, "Wireless Sensing System for Long-Time Assistance in the Parkinson's Disease," *Proceedings*, vol. 1, no. 4, p. 565, 2017. [Online]. Available: http://www.mdpi.com/2504-3900/1/4/565

[21] A. Nieuwboer, L. Rochester, T. Herman, W. Vandenberghe, G. E. Emil, T. Thomaes, and N. Giladi, "Reliability of the new freezing of gait questionnaire: Agreement between patients with Parkinson's disease and their carers," *Gait and Posture*, vol. 30, pp. 459–463, 2009.

[22] World Health Organization, "Neurological disorders: a public health approach," *Neurological disorders: public health challenges.*, pp. 41–176, 2006.

[23] E. C. Wolters, "Variability in the clinical expression of Parkinson's disease," *Journal of the Neurological Sciences*, vol. 266, no. 1-2, pp. 197–203, 2008.

[24] L. C. Triarhou, "Dopamine and Parkinson's Disease," 2013. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK6271/

[25] T. Pringsheim, N. Jette, A. Frolkis, and T. D. L. Steeves, "The prevalence of Parkinson's disease: A systematic review and meta-analysis," *Movement Disorders*, vol. 29, no. 13, pp. 1583–1590, 2014.

[26] "Tremor in Parkinson's | APDA." [Online]. Available: https://www.apdaparkinson.org/what-is-parkinsons/symptoms/tremor/

[27] R. Ashour and J. Jankovic, "Joint and skeletal deformities in Parkinson's disease, multiple system atrophy, and progressive supranuclear palsy," *Movement Disorders*, vol. 21, no. 11, pp. 1856–1863, nov 2006. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/16941460http://doi.wiley.com/10.1002/mds.21058

[28] A. Berardelli, J. C. Rothwell, P. D. Thompson, M. Hallett, and A. Berardelli, "Pathophysiology of bradykinesia in Parkinson's disease," *Brain*, vol. 124, pp. 2131–2146, 2001.

[29] S. Fahn, "Movement Disorders: Overview," in *Encyclopedia of Movement Disorders*. Elsevier, 2010, pp. 209–219. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/B9780123741059003531

[30] "Common Symptoms of Parkinson's Disease | American Parkinson Disease Assoc." [Online]. Available: https://www.apdaparkinson.org/what-is-parkinsons/symptoms/{#}motor

[31] N. Giladi, T. A. Treves, E. S. Simon, H. Shabtai, Y. Orlov, B. Kandinov, D. Paleacu, and A. D. Korczyn, "Freezing of gait in patients with advanced Parkinson's disease," *Journal of Neural Transmission*, vol. 108, pp. 53–61, 2001.

[32] M. Macht, Y. Kaussner, J. C. Möller, K. Stiasny-Kolster, K. M. Eggert, H. P. Krüger, and H. Ellgring, "Predictors of freezing in Parkinson's disease: A survey of 6,620 patients," *Movement Disorders*, vol. 22, no. 7, pp. 953–956, 2007.

[33] N. Giladi and A. Nieuwboer, "Understanding and treating freezing of gait in Parkinsonism, proposed working definition, and setting the stage," *Movement Disorders*, vol. 23, no. SUPPL. 2, pp. 423–425, 2008.

[34] S. Fahn, "The freezing phenomenon in parkinsonism." *Advances in neurology*, vol. 67, pp. 53–63, 1995. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/8848982

[35] N. Giladi, J. Tal, T. Azulay, O. Rascol, D. J. Brooks, E. Melamed, W. Oertel, W. H. Poewe, F. Stocchi, and E. Tolosa, "Validation of the Freezing of Gait Questionnaire in patients with Parkinson's disease," *Movement Disorders*, vol. 24, pp. 655–661, 2009.

[36] J. Spildooren, S. Vercruysse, K. Desloovere, W. Vandenberghe, E. Kerckhofs, and A. Nieuwboer, "Freezing of Gait in Parkinson's Disease: The Impact of Dual-Tasking and Turning," *Movement Disorders*, vol. 25, pp. 2563–2570, 2010. [Online]. Available: https://lirias.kuleuven.be/bitstream/123456789/273685/3/Spildooren{_}et{_}al-2010-Freezing+of+gait+in+Parkinson{%}27s+disease+\T1\textendash+the+impact+of+dual-tasking+and+turning.pdf

[37] M. Plotnik, N. Giladi, and J. M. Hausdorff, "Is freezing of gait in Parkinson's disease a result of multiple gait impairments? Implications for treatment," *Parkinson's Disease*, vol. 2012, 2012.

[38] S. J. G. Lewis and R. A. Barker, "A pathophysiological model of freezing of gait in Parkinson's disease," *Parkinsonism and Related Disorders*, vol. 15, pp. 333–338, 2009. [Online]. Available: https://ac.els-cdn.com/S1353802008002514/1-s2.0-S1353802008002514-main.pdf?{_}tid=d9222692-f661-11e7-a851-00000aacb35d{&}acdnat=1515628767{_}2908c539073e75a978c4e03f9764e0e3

[39] J. Vandenbossche, N. Deroost, E. Soetens, D. Coomans, J. Spildooren, S. Vercruysse, A. Nieuwboer, E. Kerckhofs, S. S. Nagarajan, P. Starr, and K. Ganguly, "Freezing of gait in Parkinson's disease: disturbances in automaticity and control," *Frontiers in Human Neuroscience*, vol. 6, p. 356, 2013. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3541536/pdf/fnhum-06-00356.pdf

[40] J. V. Jacobs, J. G. Nutt, P. Carlson-Kuhta, M. Stephens, and F. B. Horak, "Knee trembling during freezing of gait represents multiple anticipatory postural adjustments," *Experimental Neurology*, vol. 215, no. 2, pp. 334–341, feb 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0014488608004147

[41] A. Nieuwboer and N. Giladi, "Characterizing freezing of gait in Parkinson's disease: Models of an episodic phenomenon," *Movement Disorders*, vol. 28, no. 11, 2013.

[42] K. Marder, M. X. Tang, H. Mejia, B. Alfaro, L. Côté, E. Louis, J. Groves, and R. Mayeux, "Risk of Parkinson's disease among first-degree relatives: A community-based study." *Neurology*, vol. 47, no. 1, pp. 155–60, jul 1996. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/8710070

[43] I. Arnulf, E. Konofal, M. Merino–Andreu, J. L. Houeto, V. Mesnage, M. L. Welter, L. Lacomblez, J. L. Golmard, J. P. Derenne, and Y. Agid, "Parkinson's disease and sleepiness," *Neurology*, vol. 58, no. 7, pp. 1019 LP – 1024, apr 2002. [Online]. Available: http://n.neurology.org/content/58/7/1019.abstract

[44] Y. J. Zhao, H. L. Wee, Y. H. Chan, S. H. Seah, W. L. Au, P. N. Lau, E. C. Pica, S. C. Li, N. Luo, and L. C. S. Tan, "Progression of Parkinson's disease as evaluated by Hoehn and Yahr stage transition times," *Movement Disorders*, vol. 25, no. 6, pp. 710–716, 2010.

[45] C. G. Goetz, W. Poewe, O. Rascol, C. Sampaio, G. T. Stebbins, C. Counsell, N. Giladi, R. G. Holloway, C. G. Moore, G. K. Wenning, M. D. Yahr, and L. Seidl, "Movement Disorder Society Task Force Report on the Hoehn and Yahr Staging Scale: Status and Recommendations," *Movement Disorders*, vol. 19, no. 9, pp. 1020–1028, 2004. [Online]. Available: http://display.mds.prod.titanclient.com/MDS-Files1/PDFs/Task-Force-Papers/hoehnyahr.pdf

[46] A. H. Snijders, C. A. Haaxma, Y. J. Hagen, M. Munneke, and B. R. Bloem, "Freezer or non-freezer: Clinical assessment of freezing of gait," *Parkinsonism and Related Disorders*, vol. 18, no. 2, pp. 149–154, 2012.

[47] J. Nonnekes, A. M. Janssen, S. H. G. Mensink, L. B. Oude Nijhuis, B. R. Bloem, and A. H. Snijders, "Short rapid steps to provoke freezing of gait in Parkinson's disease," *Journal of Neurology*, vol. 261, no. 9, pp. 1763–1767, 2014.

[48] J. D. Schaafsma, Y. Balash, T. Gurevich, A. L. Bartels, J. M. Hausdorff, and N. Giladi, "Characterization of freezing of gait subtypes and the response of each to levodopa in Parkinson's disease," *European Journal of Neurology*, vol. 10, no. 4, pp. 391–398, 2003.

[49] K. Ziegler, F. Schroeteler, A. O. Ceballos-Baumann, and U. M. Fietzek, "A new rating instrument to assess festination and freezing gait in Parkinsonian patients," *Movement Disorders*, vol. 25, no. 8, pp. 1012–1018, 2010.

[50] E. Barry, R. Galvin, C. Keogh, F. Horgan, and T. Fahey, "Is the Timed Up and Go test a useful predictor of risk of falls in community dwelling older adults: a systematic review and meta- analysis," *BMC Geriatrics*, vol. 14, pp. 1–14, 2014. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3924230/pdf/1471-2318-14-14.pdf

[51] N. Giladi, "Medical treatment of freezing of gait," *Movement Disorders*, vol. 23, pp. 482–488, 2008.

[52] O. Dehzangi, M. Taherisadr, and R. ChangalVala, "IMU-Based Gait Recognition Using Convolutional Neural Networks and Multi-Sensor Fusion," *Sensors*, vol. 17, no. 12, pp. 1–22, 2017. [Online]. Available: http://www.mdpi.com/1424-8220/17/12/2735

[53] A. Özdemir and B. Barshan, "Detecting Falls with Wearable Sensors Using Machine Learning Techniques," *Sensors*, vol. 14, no. 6, pp. 10 691–10 708, 2014. [Online]. Available: http://www.mdpi.com/1424-8220/14/6/10691/

[54] E. Bergamini, G. Ligorio, A. Summa, G. Vannozzi, A. Cappozzo, and A. M. Sabatini, "Estimating orientation using magnetic and inertial sensors and different sensor fusion approaches: Accuracy assessment in manual and locomotion tasks," *Sensors (Switzerland)*, vol. 14, no. 10, pp. 18 625–18 649, 2014.

[55] S. Ayub, A. Bahraminisaab, and B. Honary, "A Sensor Fusion Method for Smart phone Orientation Estimation," *13th Annual Post Graduate Symposium on the Convergence of Telecommunications, Networking and Broadcasting*, 2012. [Online]. Available: http://www.cms.livjm.ac.uk/pgnet2012/Proceedings/Papers/1569603133.pdf

[56] B. Mccarron, "Low-Cost IMU Implementation via Sensor Fusion Algorithms in the Arduino Environment," Tech. Rep., 2013. [Online]. Available: http://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?article=1114{&}context=aerosp

[57] S. T. Moore, D. A. Yungher, T. R. Morris, V. Dilda, H. G. MacDougall, J. M. Shine, S. L. Naismith, and S. J. Lewis, "Autonomous identification of freezing of gait in Parkinson's disease from lower-body segmental accelerometry," *Journal of NeuroEngineering and Rehabilitation*, vol. 10, p. 1, 2013. [Online]. Available: http://www.jneuroengrehab.com/content/10/1/19

[58] J. M. Hausdorr, Y. Balash, and N. Giladi, "Time series analysis of leg movements during freezing of gait in Parkinson's disease: akinesia, rhyme or reason?" *Physica A*, vol. 321, pp. 565–570, 2003. [Online]. Available: www.elsevier.com/locate/physa

[59] A. M. A. Handojoseno, J. M. Shine, T. N. Nguyen, Y. Tran, S. J. G. Lewis, and H. T. Nguyen, "The detection of Freezing of Gait in Parkinson's disease patients using EEG signals based on Wavelet decomposition," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, aug 2012,

pp. 69–72. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/23365834http://ieeexplore.ieee.org/document/6345873/

[60] S. Mazilu, A. Calatroni, E. Gazit, A. Mirelman, J. M. Hausdorff, and G. Tr, "Prediction of Freezing of Gait in Parkinson ' s From Physiological Wearables : An Exploratory Study," *IEEE J Biomed Health Inform*, vol. 19, no. 6, pp. 1843–1854, 2015.

[61] Jong Hee Han, Won Jin Lee, Tae Beom Ahn, Beom Suk Jeon, and Kwang Suk Park, "Gait analysis for freezing detection in patients with movement disorder using three dimensional acceleration system," in *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439)*. IEEE, 2003, pp. 1863–1865. [Online]. Available: http://ieeexplore.ieee.org/document/1279781/

[62] S. T. Moore, H. G. MacDougall, and W. G. Ondo, "Ambulatory monitoring of freezing of gait in Parkinson's disease," *Journal of Neuroscience Methods*, vol. 167, no. 2, pp. 340–348, 2008.

[63] M. Bächlin, D. Roggen, M. Plotnik, J. M. Hausdorff, N. Giladi, and G. Tröster, "Online Detection of Freezing of Gait in Parkinson's Disease Patients: A Performance Characterization," *Proceedings of the 4th International ICST Conference on Body Area Networks*, p. 11, 2009. [Online]. Available: http://eudl.eu/doi/10.4108/ICST.BODYNETS2009.5852

[64] S. Mazilu, M. Hardegger, Z. Zhu, D. Roggen, G. Tröster, M. Plotnik, J. M. Hausdorff, and G. Troester, "Online Detection of Freezing of Gait with Smartphones and Machine Learning Techniques," *Proceedings of the 6th International ICST Conference on Pervasive Computing Technologies for Healthcare*, no. 3, pp. 123–130, 2012. [Online]. Available: http://eudl.eu/doi/10.4108/icst.pervasivehealth.2012.248680

[65] C. Ahlrichs, A. Samà, M. Lawo, J. Cabestany, D. Rodríguez, M. Carlos, P. López, D. Sweeney, L. R. Quinlan, G. Ò. Laighin, T. Counihan, P. Browne, L. Hadas, G. Vainstein, A. Costa, R. Annicchiarico, S. Alcaine, B. Mestre, P. Quispe, À. Bayes, and A. Rodríguez, "Detecting freezing of gait with a tri - axial accelerometer in Parkinson ' s disease patients," *Medical & Biological Engineering & Computing*, 2015.

[66] P. Lorenzi, F. Irrera, R. Rao, G. Romano, and A. Kita, "Mobile Devices For The Real Time Detection Of Specific Human Motion Disorders," *IEEE Sensors Journal*, vol. 16, no. 23, pp. 8220–8227, 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7409917/

[67] E. Jovanov, E. Wang, L. Verhagen, M. Fredrickson, and R. Fratangelo, "deFOG — A real time system for detection and unfreezing of gait of Parkinson's patients," in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, sep 2009, pp. 5151–5154. [Online]. Available: http://ieeexplore.ieee.org/document/5334257/

[68] A. Delval, A. H. Snijders, V. Weerdesteyn, J. E. Duysens, L. Defebvre, N. Giladi, and B. R. Bloem, "Objective detection of subtle freezing of gait episodes in Parkinson's disease," *Movement Disorders*, vol. 25, no. 11, pp. 1684–1693, aug 2010. [Online]. Available: http://doi.wiley.com/10.1002/mds.23159

[69] S. J. Preece, J. Y. Goulermas, L. P. J. Kenney, D. Howard, K. Meijer, and R. Crompton, "Activity identification using body-mounted sensors—a review of classification techniques," *Physiological Measurement*, vol. 30, no. 4, p. R1, apr 2009. [Online]. Available: http://iopscience.iop.org/article/10.1088/0967-3334/30/4/R01/meta

[70] A. Nieuwboer, R. Dom, W. De Weerdt, K. Desloovere, S. Fieuws, and E. Broens-Kaucsik, "Abnormalities of the Spatiotemporal Characteristics of Gait at the Onset of Freezing in Parkinson's Disease," *Movement Disorders*, vol. 16, no. 6, pp. 1066–1075, 2001. [Online]. Available: https://lirias.kuleuven.be/bitstream/123456789/19556/1/Nieuwboer{_}et{_}al-2001-Abnormalities+of+the+spatiotemporal+characteristics+of+gait+at+the+onset+of+freezing+in+Parkinson{%}27s+disease..pdf

[71] M. L. Ferster, S. Mazilu, and G. Tröster, "Gait Parameters Change Prior to Freezing in Parkinson's Disease: A Data-Driven Study with Wearable Inertial Units," *Proceedings of the 10th EAI International Conference on Body Area Networks*, 2015. [Online]. Available: http://eudl.eu/doi/10.4108/eai.28-9-2015.2261411

[72] "The PKG™ System." [Online]. Available: https://www.globalkineticscorporation.com.au/the-pkg-system/

[73] R. I. Griffiths, K. Kotschet, S. Arfon, Z. Ming Xu, W. Johnson, J. Drago, A. Evans, P. Kempster, S. Raghav, and M. K. Horne, "Automated Assessment of Bradykinesia and Dyskinesia in Parkinson's Disease," *Journal of Parkinson's Disease*, vol. 2, pp. 47–55, 2012. [Online]. Available: https://content.iospress.com/download/journal-of-parkinsons-disease/jpd11071?id=journal-of-parkinsons-disease{%}2Fjpd11071

[74] "Kinesia ONE™ Product Overview - Kinesia." [Online]. Available: https://glneurotech.com/kinesia/products/kinesia-one/

[75] "Kinesia 360™ Product Overview - Kinesia." [Online]. Available: http://glneurotech.com/kinesia/products/kinesia-360/

[76] C. L. Vaughan, "Theories of bipedal walking: an odyssey." *Journal of biomechanics*, vol. 36, no. 4, pp. 513–23, apr 2003. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/12600342

[77] J. Liu, T. Kanno, M. Akashi, W. Chen, D. Wei, G. Wu, and N. Takeda, "Patterns of Bipedal Walking on Tri-axial Acceleration Signals and Their Use in Identifying Falling Risk of Older People," in *The Sixth IEEE International Conference on Computer*

and Information Technology (CIT'06). IEEE, 2006, pp. 205–205. [Online]. Available: http://ieeexplore.ieee.org/document/4019976/

[78] W. Tao, T. Liu, R. Zheng, and H. Feng, "Gait analysis using wearable sensors." *Sensors (Basel, Switzerland)*, vol. 12, no. 2, pp. 2255–83, 2012. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/22438763http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3304165

[79] A. Kharb, V. Saini, Y. K. Jain, and S. Dhiman, "A review of gait cycle and its parameters," *IJCEM International Journal of Computational Engineering & Management ISSN*, vol. 13, pp. 2230–7893, 2011. [Online]. Available: www.IJCEM.orgwww.ijcem.org

[80] J. A. Lee, S. H. Cho, Y. J. Lee, H. K. Yang, and J. W. Lee, "Portable activity monitoring system for temporal parameters of gait cycles," *Journal of Medical Systems*, vol. 34, pp. 959–966, 2010.

[81] "Welcome to Python.org." [Online]. Available: https://www.python.org/

[82] "PyCharm: Python IDE for Professional Developers by JetBrains." [Online]. Available: https://www.jetbrains.com/pycharm/

[83] T. T. Pham, S. T. Moore, S. J. Lewis, D. N. Nguyen, E. Dutkiewicz, A. J. Fuglevand, A. L. McEwan, and P. H. Leong, "Freezing of Gait Detection in Parkinson's Disease: A Subject-Independent Detector Using Anomaly Scores," *IEEE Transactions on Biomedical Engineering*, pp. 1–1, 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7845616/

[84] "scipy.signal.filtfilt — SciPy v0.14.0 Reference Guide." [Online]. Available: https://docs.scipy.org/doc/scipy-0.14.0/reference/generated/scipy.signal.filtfilt.html

[85] M. Duarte, "Detection of peaks in data." [Online]. Available: http://nbviewer.jupyter.org/github/demotu/BMC/blob/master/notebooks/DetectPeaks.ipynb

[86] M. Zhang and A. A. Sawchuk, "A Feature Selection-Based Framework for Human Activity Recognition Using Wearable Multimodal Sensors," *Proceedings of the 6th International Conference on Body Area Networks*, pp. 92—-98, 2011.

[87] F. Petitjean, A. Ketterlin, and P. Gançarski, "A global averaging method for dynamic time warping, with applications to clustering," *Pattern Recognition*, 2011.

[88] C. Xu, J. He, X. Zhang, C. Wang, and S. Duan, "Detection of Freezing of Gait Using Template-Matching-Based Approaches," *Journal of Sensors*, p. 8, 2017.

[89] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 26, no. 1, pp. 43–49, feb 1978. [Online]. Available: http://ieeexplore.ieee.org/document/1163055/

[90] K. Vasimalla, N. Challa, and S. Manohar Naik, "Efficient dynamic time warping for time series classification," *Indian Journal of Science and Technology*, 2016.

[91] S. Salvador and P. Chan, "FastDTW: Toward Accurate Dynamic Time Warping in Linear Time and Space," *Intelligent Data Analysis*, vol. 11, no. 5, pp. 561—-580, 2007. [Online]. Available: https://pdfs.semanticscholar.org/05a2/0cde15e172fc82f32774dd0cf4fe5827cad2.pdf

[92] V. Niennattrakul, D. Srisai, and C. A. Ratanamahatana, "Shape-based template matching for time series data," *Knowledge-Based Systems*, vol. 26, pp. 1–8, 2012.

[93] V. Niennattrakul and C. A. Ratanamahatana, "Inaccuracies of Shape Averaging Method Using Dynamic Time Warping for Time Series Data," *Proceedings of the 7th International Conference on Computational Science (ICCS'07)*, pp. 513–520, 2007. [Online]. Available: http://link.springer.com/10.1007/978-3-540-72584-8{_}68

[94] "tslearn's documentation — tslearn 0.1.17 documentation," 2017. [Online]. Available: http://tslearn.readthedocs.io/en/latest/index.html{#}

[95] S. Bersch, D. Azzi, R. Khusainov, I. Achumba, and J. Ries, "Sensor Data Acquisition and Processing Parameters for Human Activity Classification," *Sensors*, vol. 14, no. 3, pp. 4239–4270, 2014. [Online]. Available: http://www.mdpi.com/1424-8220/14/3/4239/

[96] J. B. Kinney and G. S. Atwal, "Equitability, mutual information, and the maximal information coefficient," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, pp. 3354–3359, 2014. [Online]. Available: http://www.pnas.org/content/pnas/111/9/3354.full.pdf

[97] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Computers & Electrical Engineering*, vol. 40, no. 1, pp. 16–28, jan 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045790613003066

[98] S. Raschka, "MLxtend: Providing machine learning and data science utilities and extensions to Python's scientific computing stack," *The Journal of Open Source Software*, vol. 3, no. 24, 2018. [Online]. Available: http://joss.theoj.org/papers/10.21105/joss.00638

[99] Sebastian Raschka, "SequentialFeatureSelector - mlxtend," 2018. [Online]. Available: https://rasbt.github.io/mlxtend/user{_}guide/feature{_}selection/SequentialFeatureSelector/

[100] J. VanderPlas, *Python Data Science Handbook*. O'Reilly Media, 2016. [Online]. Available: http://shop.oreilly.com/product/0636920034919.do

[101] V. Cortes, Corinna and Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273—-297, 1995.

[102] "RBF SVM parameters — scikit-learn 0.19.1 documentation." [Online]. Available: http://scikit-learn.org/stable/auto{_}examples/svm/plot{_}rbf{_}parameters.html

[103] E. E. Tripoliti, A. T. Tzallas, M. G. Tsipouras, G. Rigas, P. Bougia, M. Leontiou, S. Konitsiotis, M. Chondrogiorgi, S. Tsouli, and D. I. Fotiadis, "Automatic detection of freezing of gait events in patients with Parkinson's disease," *Computer Methods and Programs in Biomedicine*, 2013. [Online]. Available: https://ac.els-cdn.com/S0169260712002787/1-s2.0-S0169260712002787-main.pdf?{_}tid=3ccfeb1a-132d-11e8-ab71-00000aacb362{&}acdnat=1518794745{_}424612c198e0ad0a731952525635315c

[104] A. Mannini and A. M. Sabatini, "Machine Learning Methods for Classifying Human Physical Activity from On-Body Accelerometers," *Sensors*, vol. 10, pp. 1154–1175, 2010.

[105] T. K. Kim, "T test as a parametric statistic," *Korean Journal of Anesthesiology*, 2015. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4667138/pdf/kjae-68-540.pdf

[106] Y. Benjamini and D. Yekutieli, "The control of the false discovery rate in multiple testing under dependency," *The Annals of Statistics*, vol. 29, no. 4, pp. 1165–1188, 2001. [Online]. Available: https://projecteuclid.org/download/pdf{_}1/euclid.aos/1013699998

[107] S. Belia, F. Fidler, J. Williams, and G. Cumming, "Researchers Misunderstand Confidence Intervals and Standard Error Bars." *Psychological Methods*, vol. 10, no. 4, pp. 389–396, dec 2005. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/16392994http://doi.apa.org/getdoi.cfm?doi=10.1037/1082-989X.10.4.389